Hewlett Packard Enterprise

Reference Architecture

HPE Reference Configuration for Red Hat OpenShift Container Platform on HPE Synergy and HPE 3PAR StoreServ Storage

Contents

Executive summary

In today's digital world, organizations are under increasing pressure to deliver applications faster while reducing costs. As these applications grow more complex, this puts stress on IT infrastructure, IT teams, and processes. To remain competitive, organizations must adapt quickly and developers need to be more effective, efficient, and agile. Container technology provides the right application platform to help organizations become more responsive and iterate across multiple IT environments as well as develop, deploy, and manage applications faster. But implementing a containerized environment across existing infrastructure is a complex undertaking that can require weeks or months to mobilize, particularly for enterprises. To help accelerate container application delivery, Hewlett Packard Enterprise and Red Hat® are collaborating to optimize Red Hat OpenShift Container Platform on HPE platforms, including HPE Synergy, the industry's first composable infrastructure and HPE 3PAR Storage.

Red Hat OpenShift Container Platform on HPE Synergy provides an end-to-end fully integrated container solution that, once assembled, can be configured within hours. This eliminates the complexities associated with implementing a container platform across an enterprise data center and provides the automation of hardware and software configuration to quickly provision and deploy a containerized environment at scale. Red Hat OpenShift Container Platform provides organizations with a reliable platform for deploying and scaling container-based applications. HPE Synergy provides the flexible infrastructure you need to run that container platform to dynamically provision and scale applications, whether they run in VMs or containers, or are hosted on-premises, in the cloud, or somewhere in between.

This Reference Configuration provides architectural guidance for deploying, scaling and managing a Red Hat OpenShift environment on Hewlett Packard Enterprise Synergy Composable Infrastructure along with HPE 3PAR StoreServ Storage.

This Reference Configuration describes how to:

- Efficiently lay out an OpenShift configuration using a mix of virtual machines and bare metal hosts.
- Configure persistent storage for containers using HPE 3PAR StoreServ Storage.

This Reference Configuration demonstrates the following benefits of utilizing HPE Synergy for Red Hat OpenShift Container Platform:

- Automated initial installation and configuration of highly available vSphere hosts and the management virtual machine for the Red Hat OpenShift Container Platform, is reduced from more than 3 hours to approximately one hour, and complexity of the manual operation is reduced from needing to perform more than 500 steps, to running two Ansible play books.
- Automated deployment and configuration of physical worker nodes is reduced from 8 hours to under 20 minutes, and complexity of manual operation is reduced from needing to perform close to 300 steps, to running two Ansible play books.
- Automated host preparation of the Red Hat OpenShift Container Platform Nodes is reduced from up to 8 hours to as little as 20 minutes, and the complexity of the manual operation is reduced from performing more than three hundred steps, to one Ansible playbook.
- Deployment of the core management functions on VMs to optimize resource usage while keeping the worker nodes on bare metal to optimize performance.
- Disconnected installation: It ensures the OpenShift Container Platform software is made available to the relevant servers, then follows the same installation process as a standard connected installation.
- Using an enterprise grade storage solution such as HPE 3PAR StoreServ Storage for persistent storage, for containers enables speed, portability, and agility for traditional enterprise applications and data.
- The HPE Composable Infrastructure solution provides a layered view of security controls. The objective of choosing this layered security view is to ensure that consumers become aware of the depth of security risk that an infrastructure can have, and also to make them aware of the depth of defense that is built in to the HPE Synergy Composable Infrastructure design.
- Container Platform security provided by Sysdig, that is installed as part Sysdig cloud-native intelligence platform, offers unified container security, monitoring, and forensics for container and Kubernetes based environments.
- A business-driven container application data protection architecture provided by HPE Recovery Manager Central (RMC) and HPE StoreOnce.
- Value add services like service mesh, cluster console, Prometheus monitoring, and Grafana dashboard are also provided.

Target audience: This document is for Chief Information Officers (CIOs), Chief Technology Officers (CTOs), data center managers, enterprise architects and implementation personnel wishing to learn more about Red Hat OpenShift Container Platform on HPE Synergy Composable Infrastructure. Familiarity with HPE Synergy, HPE 3PAR StoreServ storage, Red Hat OpenShift Container Platform, container-based solutions, Ansible® Engine and core networking knowledge is assumed.

Document purpose: The purpose of this document is to provide the benefits and technical details of deploying Red Hat OpenShift on HPE Synergy Composable Infrastructure with HPE 3PAR StoreServ storage. The implementation details, processes and plays discussed in this document are available from Hewlett Packard Enterprise at https://github.com/hewlettpackard/hpe-solutions-openshift.

Introduction

This Reference Configuration describes a highly available and secure Red Hat OpenShift Container Platform deployment on HPE Synergy Composable Infrastructure, including details on how the environment is connected and configured. When combined with the accompanying Deployment Guide (https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/3par-vsphere), it provides a comprehensive example demonstrating how Red Hat OpenShift Container Platform can be set up to take advantage of the HPE Synergy Composable architecture and leverage HPE 3PAR StoreServ Storage. The configuration used for this solution consists of three (3) OpenShift Container Platform master instances and six (6) OpenShift Container Platform worker instances. This configuration is housed on a three (3) frame HPE Synergy with nine (9) HPE Synergy 480 Gen 10 Compute Modules installed with a mix of VMware® vSphere® and Red Hat® Enterprise Linux®. The HPE 3PAR StoreServ storage is used to provide persistent storage for containers and registry, virtual machine storage and data management and supports both iSCSI and Fibre Channel connectivity. This architecture can scale between three (3) to n worker nodes.

Due to the ephemeral nature of containers, protecting persistent data associated with the containers becomes a crucial task. In this Reference Configuration, the Red Hat OpenShift Container Platform pod's persistent volume is protected using HPE RMC, which initiates a crash consistent snapshot at the volume level and, using the RMC express protect feature, moves the snapshot to an HPE StoreOnce Catalyst store. In this scenario no external data mover is involved. Either HPE StoreOnce or HPE RMC acts as the data mover. This helps in reducing the cost and complexity of the solution. Optionally, HPE 3PAR StoreServ Storage can be configured for replication of the volume to a remote array which reduces the recovery point objective/recovery time objective (RPO/RTO). RPO/RTO can be further reduced with peer persistent (active/active) replication.

Security is a required component of any production IT infrastructure solution. The overall containerized ecosystem solution discussed in this document is secured using Sysdig cloud-native intelligence platform. Sysdig Secure performs container image scanning, run-time protection, and forensics to identify vulnerabilities, block threats, enforce compliance, and audit activity across enterprise cloud-native environments at scale. Sysdig Monitor is a powerful application for monitoring and managing the risk, health, and performance of your microservices.

The HPE Synergy platform is designed to bridge traditional and cloud-native applications with the implementation of HPE Synergy Composable Infrastructure. HPE Synergy Composable Infrastructure combines the use of fluid resource pools made up of compute, storage, and fabric with software-defined intelligence. Composable pools of compute, storage, and fabric can be intelligently and automatically combined to support any workload. The resource pools can be flexed to meet the needs of any business application. HPE Synergy platform provides the agility and scalability on the hardware layer to the overall Red Hat OpenShift Container Platform solution.

Solution overview

This Reference Configuration provides an overview of the Red Hat OpenShift Container Platform on HPE Synergy and HPE 3PAR StoreServ storage solution as described in greater detail at https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/3par-vsphere

In this Reference Configuration, HPE has created implementation methods that allow both Fibre Channel and iSCSI connectivity on the HPE 3PAR StoreServ Storage.

Figure 1 provides an overview of the solution components.

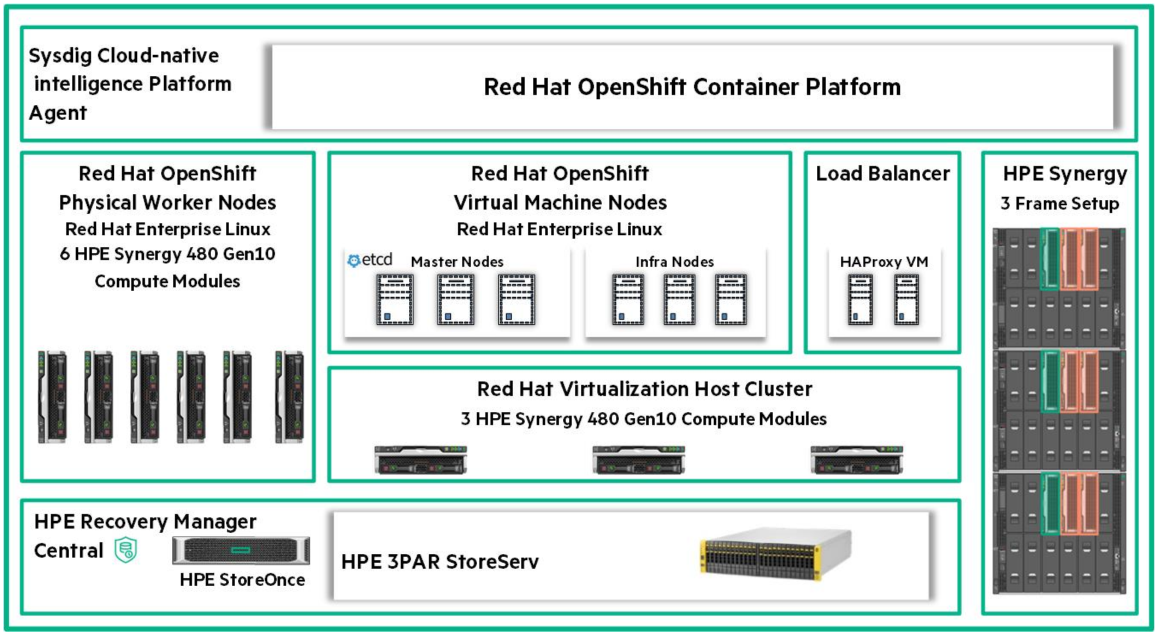


Figure 1. Solution layout

This Reference Configuration deploys Red Hat OpenShift Container Platform 3.11 as a combination of virtual and physical resources. When the SAN is connected via iSCSI, the OpenShift master and infrastructure nodes are deployed as virtual machines running on three (3) HPE Synergy 480 Gen10 Compute Modules running VMware vSphere Virtualization host, and managed by VMware vCenter Server® appliance. When connected via Fibre Channel, infrastructure functions are deployed on physical worker nodes running Red Hat Enterprise Linux version 7.6.

VMware vSphere hosts used in the solution are booted from HPE Image Streamer and are configured using Ansible playbooks. Load balancing can be deployed as a virtual machine running HAProxy or F5 BIG-IP or as physical F5 BIG-IP appliances. Red Hat OpenShift worker nodes are deployed on bare metal on six (6) HPE Synergy 480 Gen10 Compute Modules running Red Hat Enterprise Linux version 7.6 or, when the SAN is connected via iSCSI, as virtual machines. The operating system for the Red Hat OpenShift physical worker nodes is booted from HPE Image Streamer and post-installation configuration steps are performed, in part, using Ansible playbooks. Virtual machines are deployed and configured using Ansible playbooks. HPE 3PAR StoreServ storage provides support for both ephemeral and persistent container volumes.

The HPE Converged Architecture 750 (CA750) was used as the reference platform for the Red Hat OpenShift deployment from a design and firmware/software perspective. The CA750 approach provides pre-integrated, modular, scalable converged systems that reduce deployment risk. To implement this solution as an HPE CA750, please work with your Hewlett Packard Enterprise authorized channel partner. For more information about flexible HPE Converged Systems see https://www.hpe.com/us/en/integrated-systems/converged-architecture.html.

Security

To address the security challenges that exist in containerized environments, this solution leverages Sysdig SaaS Platform to secure and monitor Red Hat OpenShift Container Platform, an enterprise-ready Kubernetes platform installed and configured on HPE Synergy Composable Infrastructure. Once the configuration is deployed, access to the Red Hat OpenShift Cluster is granted to the Sysdig SaaS Platform. The Sysdig SaaS Platform is a cloud-based service where the security and monitoring services will be available to the user based on the subscription they have chosen. For security and monitoring of OpenShift Containers, it is required to install the Sysdig Agent on the OpenShift Cluster, which means Sysdig Agents that are light-weight entities will be installed within each node in the OpenShift Cluster. These agents run as a daemon to enable Sysdig Monitor and Sysdig Secure functionality. Sysdig Monitor provides deep, process-level visibility into dynamic, distributed production environments. Sysdig Secure provides image scanning, run-time protection, and forensics to identify vulnerabilities, block threats, enforce compliance, and audit activity across an OpenShift Cluster.

The key security benefits that this solution provides are:

- Faster incident resolution using Sysdig Monitor for OpenShift Cluster
- Simplified compliance for the entire solution
- Service-based access control for container security and monitoring
- Less time spent on managing platforms, containers, and vulnerabilities.

The implementation of Sysdig in this solution uses the Software as a Service (SaaS) deployment method. The playbooks deploy Sysdig Agent software on every OpenShift node and captured data is relayed back to your Sysdig SaaS Cloud portal. The deployment provides access to a 90 day try-and-buy, fully featured version of the Sysdig software.

Note
The Sysdig functionality is not turned on by default in this solution. Consult the section on Sysdig configuration for more information on how to enable Sysdig. For more information on how to access the 90 day try-and-buy version, see the GitHub repository at https://github.com/HewlettPackard/Docker-Synergy.

Figure 2 represents a security overview for Red Hat OpenShift Container Platform on HPE Synergy.

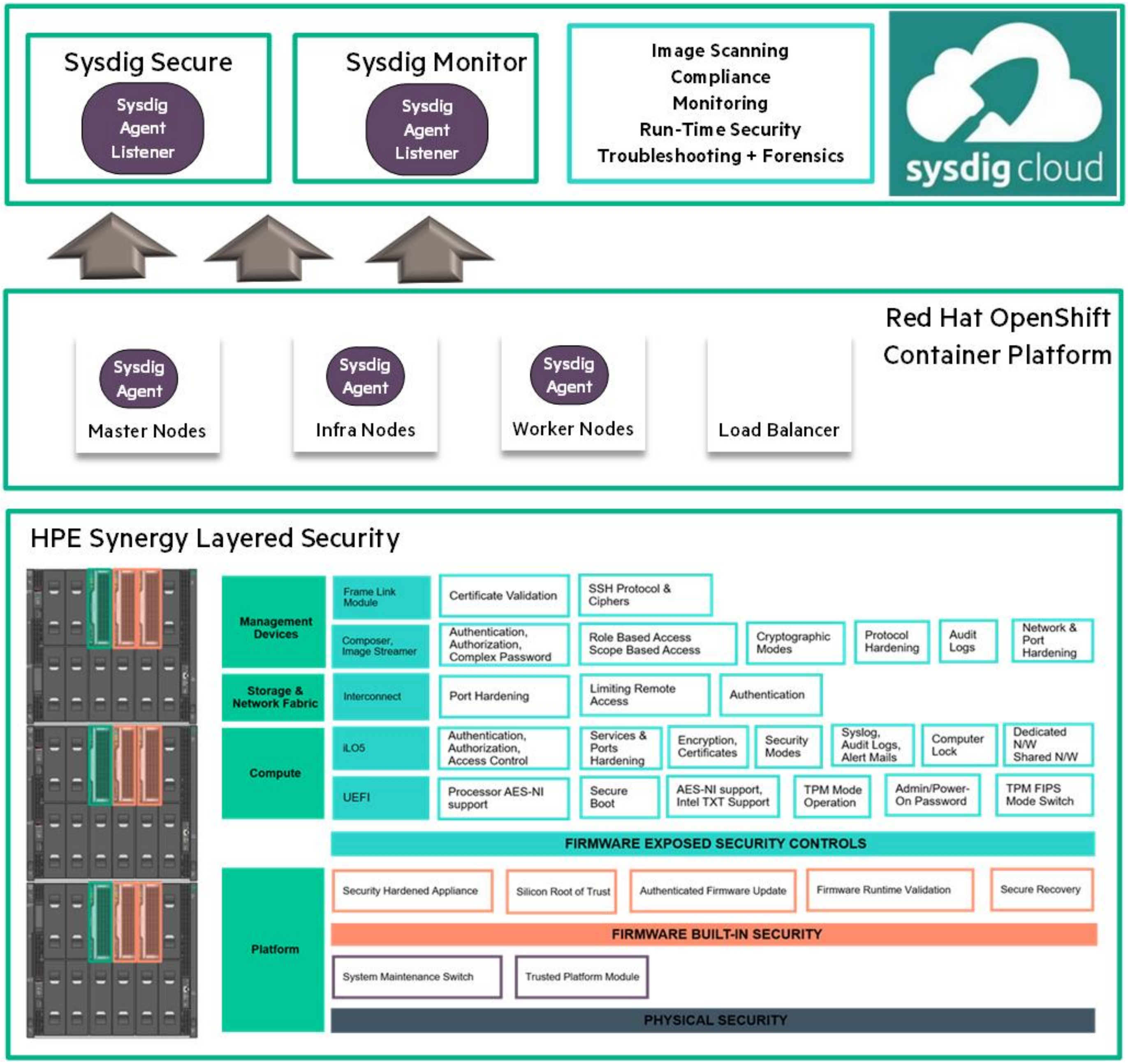


Figure 2. Security overview for Red Hat OpenShift Cluster platform on HPE Synergy platform

For more information on Sysdig Security for Red Hat OpenShift Container Platform, refer to https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/3par.

HPE Composable Infrastructure security controls

Hewlett Packard Enterprise has security features and functionalities built into servers from the hardware level to the firmware. Customers require a holistic view of the security controls available to them in the HPE Composable Infrastructure to make the most of their investment. HPE Synergy Composable Infrastructure enables IT organizations to accelerate application and service delivery through the use of fluid resource pools, made up of compute, storage, and fabric with software-defined intelligence. Each resource within the HPE Composable Infrastructure is in turn made up of multiple products such as compute modules which in turn has multiple components such as iLO, UEFI, etc. Another example is the management devices like HPE Synergy Composer, which exposes its functions using HPE OneView and the HPE Synergy Frame Link Modules. With so many products available within the HPE Composable Infrastructure, it is important to understand the security controls available within each of them and how they can be used to help avoid potential security breaches.

This solution provides a layered view of security controls that are available to Hewlett Packard Enterprise customers. Figure 2 shows the layered security view across various composable infrastructure component.

Table 1 describes the security control layers of Figure 2 - physical security controls layer, firmware/hardware built-in security controls layer and firmware exposed security functionality controls layer. The objective of choosing this layered security view is to ensure that the customers will be aware of the depth of security risk that an infrastructure can have and also to know the defense in depth that is built-in to the HPE Composable Infrastructure design. Each security control at each layer is designed to comply with the requirements of some security tenets. The security tenets are a set of security principles that ensure the security within the information systems, example: Authentication, Authorization, Access Control, Password Policies, Cryptographic Ciphers, Secure Protocols, Forensic Analytics – Logs, Alerts, Threat Modelling, Security Certifications and Standards.

Table 1: Physical and firmware based security controls within HPE Composable Infrastructure

Security controls category	Description
Physical security controls	Physical security describes measures designed to ensure the physical protection and detection of threat event in the infrastructure.
Firmware/hardware built-in security controls	This covers: • The security technologies built in the firmware to make it more secure for any communication with the underlying hardware and safe for user data at rest/transit. • The threat modelling followed within HPE to security harden the infra components.
Firmware exposed security functionality controls	This is the exhaustive list of security controls that let the customers: • Define the boundaries for accessing various infra components. • Set quantum safe ciphers for encryption. • Generate alert and log changes to infra.

Firmware built-in security controls

Hewlett Packard Enterprise has used a variety of technologies to ensure that the built-in firmware security controls provide the highest level of infrastructure security. This section provides a brief overview of the security controls that Hewlett Packard Enterprise has built into the firmware that is used by HPE Synergy and how these security controls offers an added advantage for HPE Synergy customers.

Silicon Root of Trust: The iLO5 chipset contains a first of its kind Silicon Root of Trust for the HPE Synergy Gen 10 Compute platform which is included with the iLO standard license. Silicon Root of Trust provides an inextricably tied link between the silicon and firmware—making it impossible to insert any malware, virus, or compromised code that would corrupt the boot process. The Silicon Root of Trust enables the boot process to provide a Secure Start. When the system boots, the iLO5 chip validates and boots its own firmware first, then validates the system BIOS. Because the Silicon Root of Trust is inextricably tied into the iLO5 hardware, every validated signature throughout the boot process can be trusted. However, in the unlikely event that iLO5 finds tampering or corruption at any point in the process, trusted firmware is immediately available for Secure Recovery. On startup, if iLO5 finds its own firmware that has been compromised, it will load its own authenticated firmware from an integrated backup. The iLO5 firmware recovery is always available and always automatic—regardless of license. Remember that the Silicon Root of Trust in hardware is how the iLO5 firmware is verified, so it can always be trusted. Second, if iLO5 finds that the system BIOS has been compromised, customers can connect to iLO5 and manually recover to authenticated firmware. Because the Silicon Root of Trust is embedded in the hardware itself, iLO5 is able to detect any compromised firmware—as far back as the supply chain process. Hewlett Packard Enterprise can address platform security all the way back to the supply chain because Hewlett Packard Enterprise designs the iLO5 entirely—hardware and firmware—and controls the iLO5 production process. Unlike other companies, Hewlett Packard Enterprise does not outsource the server management controller. Hewlett Packard Enterprise also has strict internal processes that dictate the firmware approval process. This gives customers an unprecedented level of assurance that no hackers have compromised the firmware before the server is received.

Secure Recovery: Secure Recovery is included in the iLO Advanced Premium Security Edition license and works alongside Silicon Root of Trust to automatically recover firmware back to a known good state in the unlikely event that it is compromised. As described previously, the Silicon Root of Trust enables the secure start process. As the system boots and iLO5 verifies the series of digital signatures, iLO5 can access trusted firmware immediately and recover to a known good state, if it finds tampering or corruption in its own firmware or the system BIOS. First, if iLO5 finds that its own firmware has been compromised, it will load its own authenticated firmware from an integrated backup. The Secure Recovery of iLO5 firmware is always available and always automatic—regardless of license. Second, if iLO5 finds that the system BIOS has been compromised, iLO5 will try to recover from a backup copy. If the backup copy is also compromised and the customer has upgraded to the iLO Advanced Premium Security Edition license, iLO5 can automatically recover authentic firmware. The standard license provides the opportunity for manual recovery. The Silicon Root of Trust is the foundation for the entire Secure State and Secure Recovery process, enabling HPE Synergy Gen10 servers to be the world's most secure industry standard servers and providing the extraordinary ability to not only verify the digital signatures up through the entire boot process but also to recover securely if any firmware is compromised.

Firmware runtime validation: With the iLO Advanced Premium Security License, the iLO5 chipset enables runtime validation of firmware. With firmware runtime verification, the iLO5 chipset performs the same checking process that happens during the boot process on a continual basis while the server is running. As frequently as once a day, iLO5—with its Silicon Root of Trust—runs a background verification check on the iLO5 firmware, UEFI, and other firmware loaded after including the SPLD, IE, and ME.

Authenticated firmware updates: The iLO5 chipset expands the number of firmware items that customers can update directly and securely in the Gen10 servers. This is a standard feature on the iLO5. Firmware items that can be securely validated and updated from the iLO now includes system programmable logic devices (SPLDs), HPE ProLiant power interface control utility (PowerPIC) firmware, the Intel® innovation Engine and Management Engine, and other low-level system components. The iLO5 contains a firmware repository stored on non-volatile flash memory (NAND), which allows components such as the Service Pack for ProLiant (SPP) and other firmware updates to be applied and installed offline through iLO5.

Best practices followed by HPE to deliver security hardened Synergy Composer appliance: Hewlett Packard Enterprise follows Secure Development Lifecycle and used a security assessment tool called Comprehensive Applications Threat Analysis (CATA) to identify and remediate security defects in the appliance operating system.

Note
The design of the appliance is based on CATA fundamentals and underwent CATA review.

The factors that contribute to appliance security hardening are listed below:

- Appliance is hardened to enforce mandatory access control. This means users of HPE Synergy are provided the role-based access control (RBAC) that enables an administrator to establish access control and authorization for users based on their responsibilities.
- Important services of the appliance run with required privileges. This implies HPE Synergy Composer is governed by scope-based access control that enables an administrator to establish access control for users by allowing a role to be restricted to a subset of resources managed by the appliance.
- The appliance is configured and maintains a firewall that blocks unused ports. Restricting the usage of all non-essential ports reduces the attack surface for HPE Synergy Composer.
- The appliance operating system bootloader is password protected. This means HPE Synergy Composer cannot be compromised by someone attempting to boot in single-user mode.
- The appliance is designed to operate in an isolated management LAN. Hewlett Packard Enterprise recommends creating a private management LAN and keeping that separate, known as air-gapped, from production LANs, using VLAN or firewall technology (or both).
- Hewlett Packard Enterprise supports digital signing of all software/firmware updates to ensure their integrity and authenticity. This implies that when the customer is re-imaging the HPE Synergy Composer in order to quickly bring it to a specific firmware revision level, the digital signature is verified by the reimaging process.

- Operating-system-level users are not allowed to access the appliance, with the following exceptions:
  - A special preset command used only if the Infrastructure administrator password is lost or forgotten. This command requires that you contact your authorized support representative to obtain a one-time password.
  - A setting that enables an authorized support representative to obtain a one-time password so that they can log in to the appliance console (and only the console) to perform advanced diagnostics. Customer can either enable or disable access with this setting.
- Hewlett Packard Enterprise closely monitors security bulletins for threats to appliance software components and, if necessary, issues software updates.

Data protection for Red Hat OpenShift

Containers have dramatically increased in popularity as organizations recognize the benefits with respect to both time and resource efficiency. This explosive growth of container applications overwhelms traditional data protection approaches. Applying traditional data protection strategies to containerized applications will not work. The goals of this solution with regards to data protection are to:

- Highlight the importance of protecting each component within an OpenShift cluster including persistent volumes in order to restore in case of corruption or system failures.
- Demonstrate Hewlett Packard Enterprise's approach to persistent volume backup using HPE Recovery Manager Central (RMC) software with the HPE 3PAR StoreServ snapshot feature and HPE StoreOnce.

OpenShift Cluster components

An OpenShift cluster is made up of several nodes and each node type has different roles. To protect the environment, it is very important to understand how these components fit together and the services provided by each component. A successful backup and recovery solution is highly unlikely without this understanding in place. This section provides details of each OpenShift node and what components require protection within the environment.

High availability is achieved using three (3) master nodes and three (3) infrastructure nodes. Worker nodes can be sized in variable quantities according to the capacity requirements of the pods that will be deployed. It is necessary to create a backup of the important components within the OCP cluster in order to recreate the nodes in the event of a failure. Figure 3 describes the major components involved in the deployment of Red Hat OpenShift Container Platform.

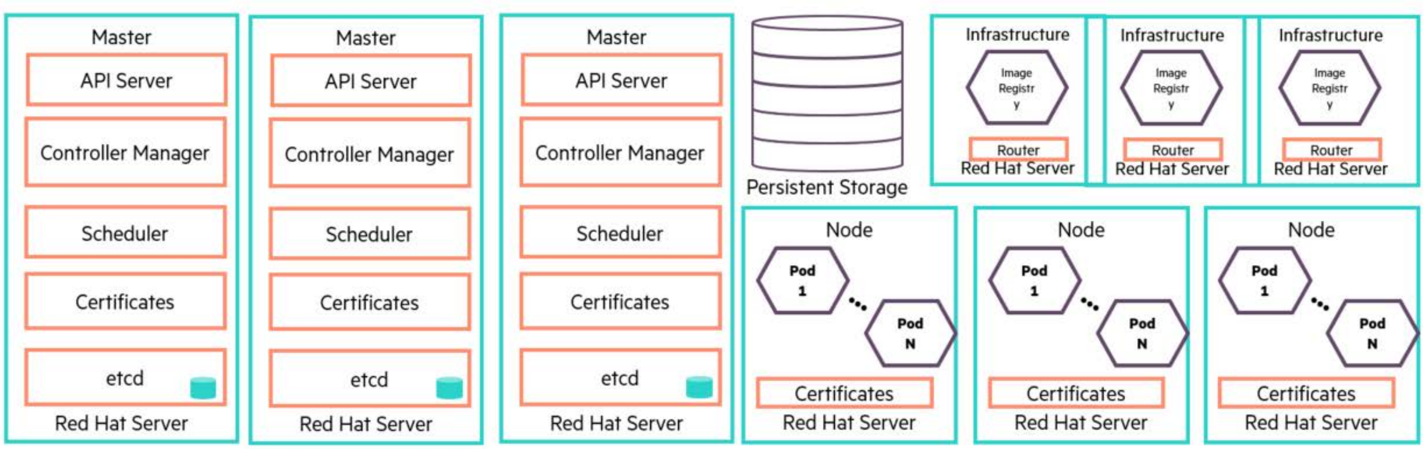


Figure 3. Components need to consider for backing up Red Hat OpenShift Container Platform solution

OpenShift master node

The OpenShift master nodes are comprised of a set of core components including:

- API server
- Controller Manager server

- Scheduler
- Certificates

The master nodes maintain the cluster's configuration, manage nodes in the OpenShift cluster, and schedule pods to run on nodes. If the OpenShift master nodes are unable to function, this will not impact the end users as the container application traffic will remain functional. However, administrators and users will not be able to make any new adjustments to the OpenShift cluster.

API server
The API server provides the management entry point of the OpenShift cluster. It mediates the interactions between the OpenShift master node components via RESTful API calls. It is responsible for storing API objects into the persistent etcd store. API server high availability is built on the persistent etcd store and deploys multiple instances of API server roles on the OpenShift cluster.

Controller Manager
The Controller Manager monitors the state of the cluster through the API Server watch feature. When a state change notification is received, it makes the necessary changes attempting to move the current state towards the desired state to keep the OpenShift cluster functioning correctly. Multiple controller manager roles are configured on OpenShift master nodes to provide high availability.

Scheduler
The scheduler ensures that container applications are scheduled to run on worker nodes within the OpenShift cluster. The scheduler reads data from the pod and attempts to find a node that is a good fit based on configured policies. To ensure high availability, more than one OpenShift master node must be configured for the scheduler roles.

Certificates
Certificates are used by the API server when securing inbound requests, authenticating users, making outbound requests, and for mutual TLS between the API server and all the other API objects in OpenShift Cluster. Certificates are copied to all the master nodes during the deployment. If more than one master host is deployed on an OpenShift cluster, the certificates are considered highly available.

etcd
etcd stores the persistent master state while other components watch etcd for changes to bring themselves into the desired state. It implements the key-value stores where all of the objects in OpenShift cluster master node components are stored. The etcd store implements a distributed consensus algorithm to ensure that even if one of the storage nodes fail, there is sufficient replication to maintain data availability. Optionally etcd role can be configured outside the master node.

OpenShift node
An OpenShift node, or worker node, provides the runtime environment for containers. Each node in an OpenShift cluster has the required services to be managed by the master. The master uses information from nodes to validate nodes with health checks. A node is ignored until it passes the health checks, and the master continues checking nodes until they are valid. Other than running pods, worker nodes contain certificates, services and authorization files. Large numbers of OpenShift nodes may be deployed in a cluster and if one node fails, it can be easily replaced without losing valuable data. However, certificates needs to be deployed on the new node. Typically, certificates and authorization files are redeployed using Ansible playbooks and the Ansible Engine/Tower will hold the files. As a result, the Ansible Engine/Tower must be protected to ensure the file is highly available.

OpenShift infrastructure node
OpenShift makes use of its local registry for storing container images. In a highly available deployment such as the one Hewlett Packard Enterprise has created, the infrastructure nodes are responsible for hosting these registry pods and this is the place where the local container images are stored. Registry pods are assigned with a Persistent Volume (PV) from external storage. In order to protect the data, it is recommended to take a snapshot or clone the volume or replicate it to a disaster recovery site.

Pods inside of an OpenShift cluster are only available via their IP addresses on the cluster network. An edge load balancer can be used to accept traffic from outside networks and route the traffic to pods inside the OpenShift cluster. An OpenShift administrator can deploy routers in an OpenShift cluster through infrastructure nodes. These enable routes created by developers to be used by external clients. OpenShift routers provide external hostname mapping and load balancing to services over protocols that pass information directly to the router. The hostname must be present in the protocol in order for the router to determine where to send traffic. For high availability, an external load balancer such as F5 BIG-IP can be used along with multiple infrastructure nodes.

Persistent storage

Containers were originally designed to run stateless applications, so there was no need for persistent storage. Once enterprises began adopting containers and they wanted to run stateful applications, persistent storage became necessary to meet the demands of the application data. HPE 3PAR StoreServ Storage provides persistent storage capabilities to an OpenShift cluster using plugins. Persistent storage and the data it houses need to be protected for business continuity, disaster recovery, and archival purpose. HPE 3PAR StoreServ integration with HPE Recovery Manager Central (RMC) and HPE StoreOnce provides snapshot protection and express protect[1] backup directly to HPE StoreOnce.

In this solution we are using Recovery Manager Central to initiate a crash consistent snapshot at the volume level and, using the RMC express protect feature, move the snapshot to the HPE StoreOnce Catalyst store. As it is moved to a backup appliance, data can be stored for archival purpose as well. In this scenario, there is no external data mover involved. Either HPE StoreOnce or HPE RMC will act as the data mover. This will reduce the cost and complexity of the solution. HPE 3PAR StoreServ Storage also supports replication of the volume to a remote array which can be used to reduce the Recovery Point Objective (RPO)/Recovery Time Objective (RTO) requirements. RPO/RTO can be further reduced with peer persistent (active/active) replication.

Figure 4 highlights the recommended data protection process for Red Hat OpenShift Container Platform using HPE 3PAR Storage.

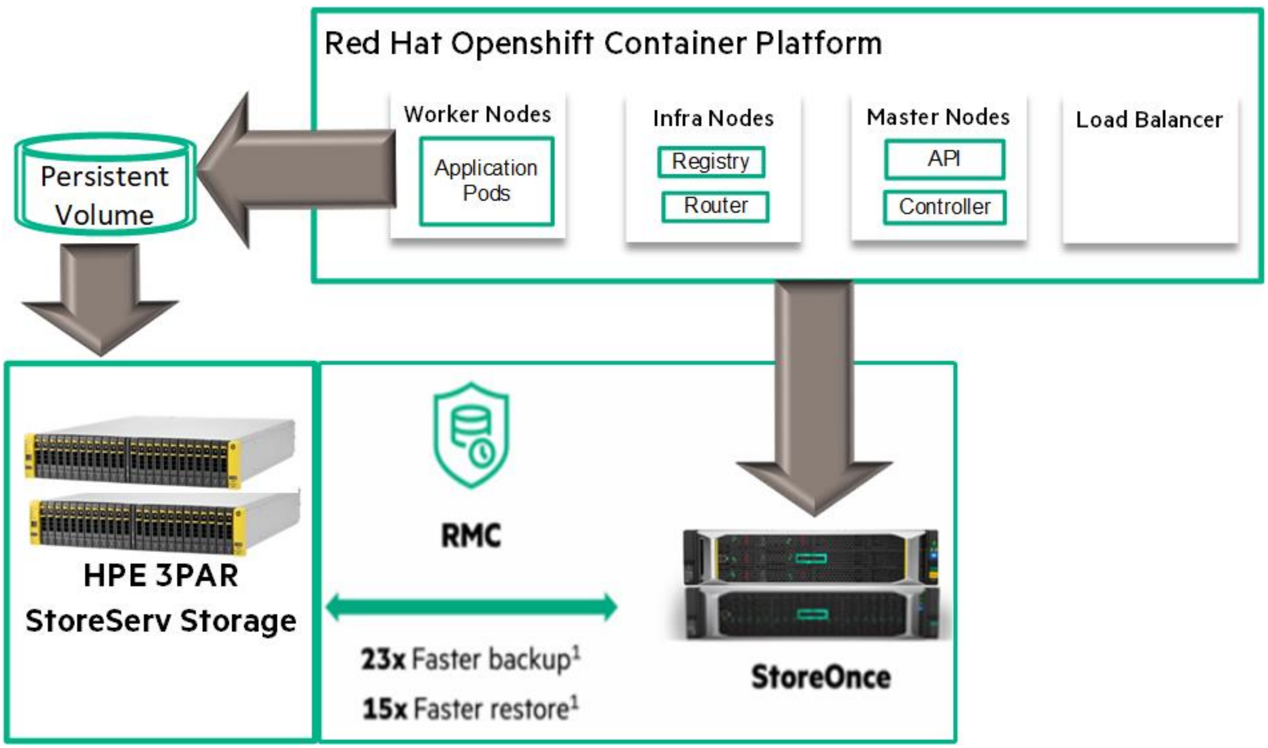


Figure 4. Data protection overview for Red Hat OpenShift Container Platform using HPE 3PAR StoreServ Storage

For more information on HPE RMC Data protection for Red Hat OpenShift Container Platform, refer to the HPE Solution Architecture for Red Hat OpenShift Container Platform Backup and Recovery on HPE Synergy and HPE 3PAR StoreServ Storage at https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable/3par-vsphere.

Node labeling in OpenShift

Discovering the node properties and advertising them through node labels can be used to control workload placement in an OpenShift cluster. But OpenShift does not have by default label nodes with any hardware configuration information. If IT wants to use hardware configuration to optimize scheduling, the capabilities of the underlying platform must be manually uncovered and labeled by administrators in order to use the

[1] RMC Express Protect Backups are faster than traditional ISV-based backups. Data is moved directly from primary storage to protection storage without any intermediate server bottlenecks.

hardware configuration in scheduling decisions. An OpenShift cluster can have many nodes and each node in turn can run multiple pods which, at scale, means that this process is both tedious and error prone. With OpenShift running on HPE server platforms, organizations can automate the discovery of hardware properties and use that information to schedule workloads that benefit from the different capabilities that the underlying hardware provides. Using HPE iLO and its REST/Redfish API based discovery capabilities (proliantutils) the following properties can be discovered about the nodes:

- Presence of GPUs
- Underlying RAID configurations
- Presence of disks by type
- Persistent-Memory availability
- Status of CPU virtualization features
- SR-IOV capabilities
- CPU architecture
- CPU core count
- Platform information including model, iLO and BIOS versions
- Memory capacity
- Status of secure boot

Once these properties are discovered for the physical worker nodes, OpenShift node labeling can be applied to group nodes based on the underlying features of the hosts. By default, every node will at least have its node name as a label. Node labels can be targeted for deployments using node selectors which can be set at either a project (can be used to restrict which nodes a project gets access to) or pod level. The node labeling system is completely open-ended, so administrator can choose whatever groupings make sense for the organization use cases.

Labels do not provide uniqueness. In general, it is expected that many objects will carry the same label(s). Using a label selector, the administrator can identify a set of objects with similar properties. This labelling can be used as either a hard or soft constraint for scheduling of application pods on desired node based on application requirements. For example, if the compute module in the HPE Composable Infrastructure has support for Intel TXT, which is specifically designed to harden platforms from the emerging threats of hypervisor attacks, malicious root kit installations, or other software-based attacks, administrators can use this information to restrict confidential data or sensitive workloads to nodes that are better controlled and have had their configurations more thoroughly evaluated through the use of Intel TXT-enabled platform.

Figure 5 shows the hardware properties that are used to label the physical worker nodes.

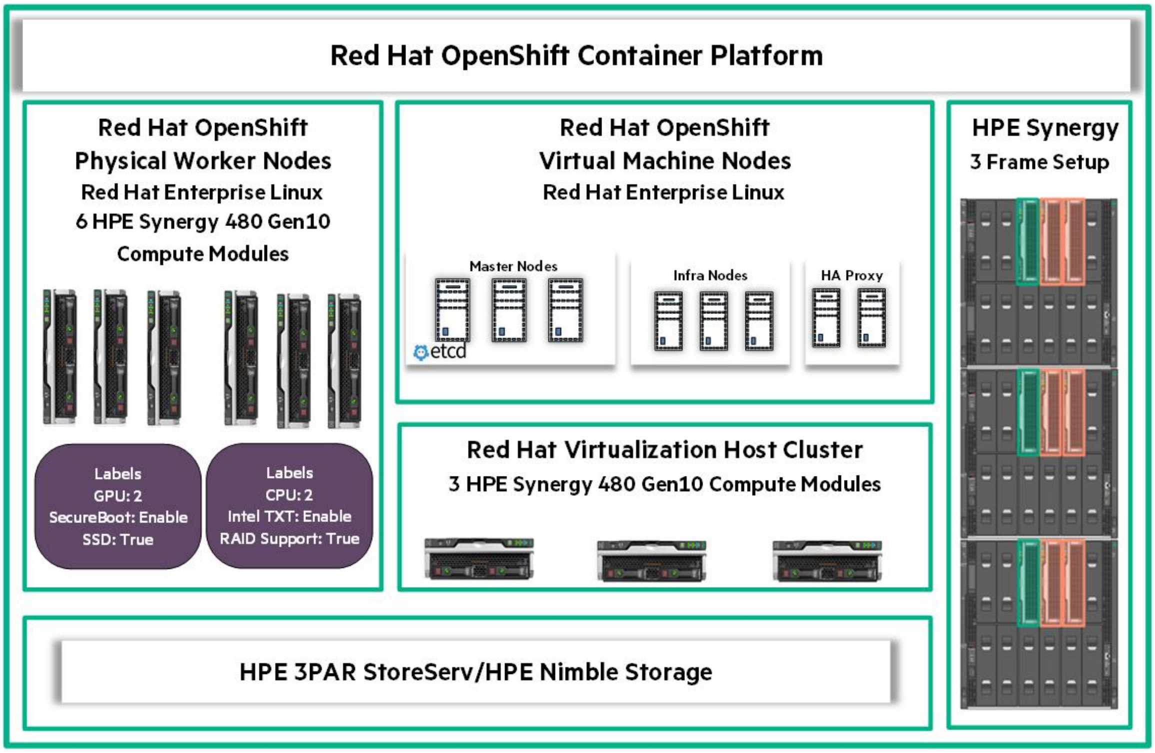


Figure 5. Physical worker node labeling

Disconnected Installation

Some data center may not have access to the internet, even via proxy servers. Installing OpenShift Container Platform in these environments is considered a disconnected installation. In these air gapped environments, OpenShift Container Platform software channels and Red Hat's Docker registry and repositories are not available via Red Hat's content distribution network. The installer will download required software repositories, packages when the installer is connected to installer's Red Hat account and will disconnect once this task is done. The installer will utilize these repositories & packages whenever they are required. The installer creates customized registry for maintaining a local registry based on web server repositories and packages. A disconnected installation ensures the OpenShift Container Platform software is made available to the relevant servers, then follows the same installation process as a standard connected installation. Follow the deployment guide at https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/3par-vsphere/iscsi/disconnected

Solution layout

Figure 6 highlights the solution at a high level. This includes a reflection of the relationship between hosts, OS/hypervisor, boot volumes and SAN storage.

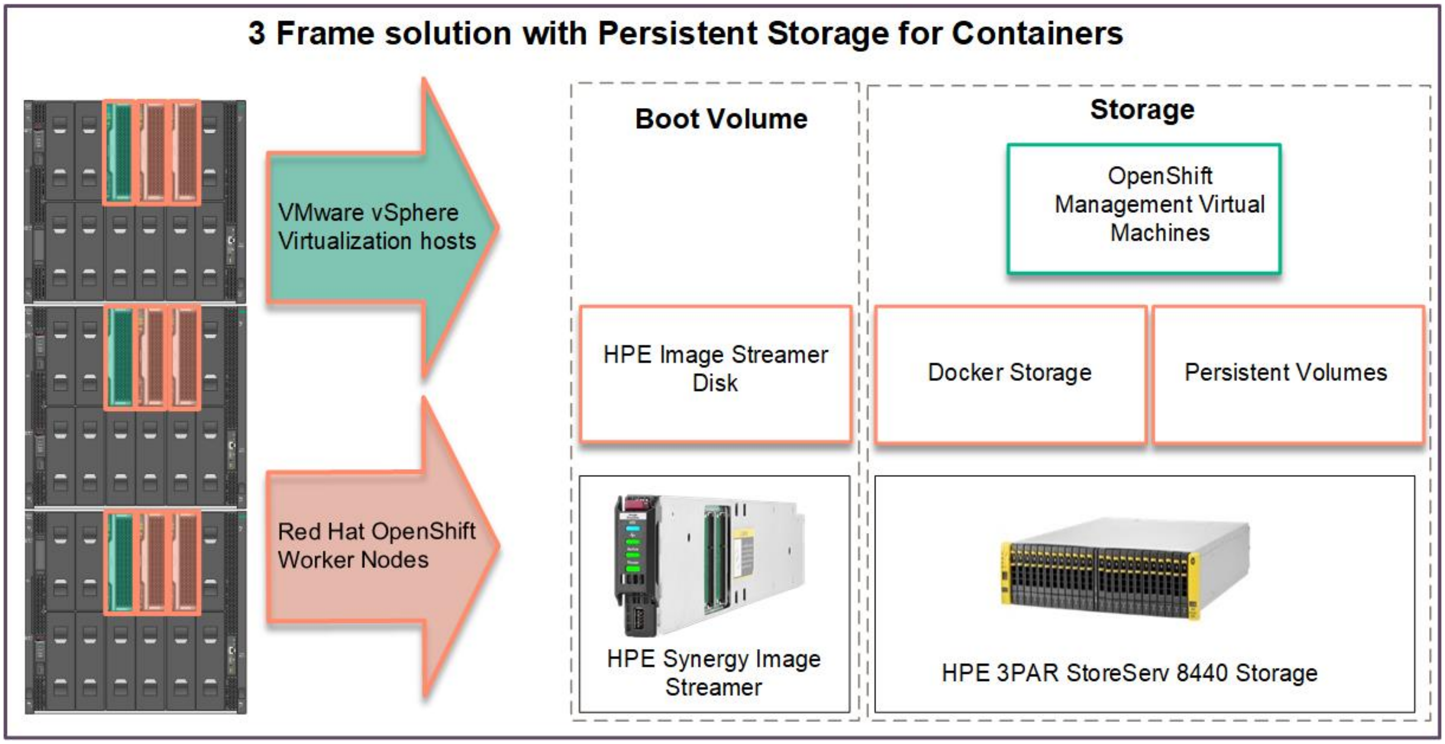


Figure 6. High level solution layout of storage resources

The solution assumes that the following infrastructure services and components are installed, configured, and function properly:

- Ansible Engine
- LDAP/Active Directory
- DHCP
- DNS
- NTP
- TFTP/PXE

Figure 7 shows the layout of the hardware within the racks. The 2 HPE SN6600B FC Switches shown are only required to support Fibre Channel connectivity to the HPE 3PAR StoreServ Storage. Implementations that utilize iSCSI for storage connectivity do not require additional switching.

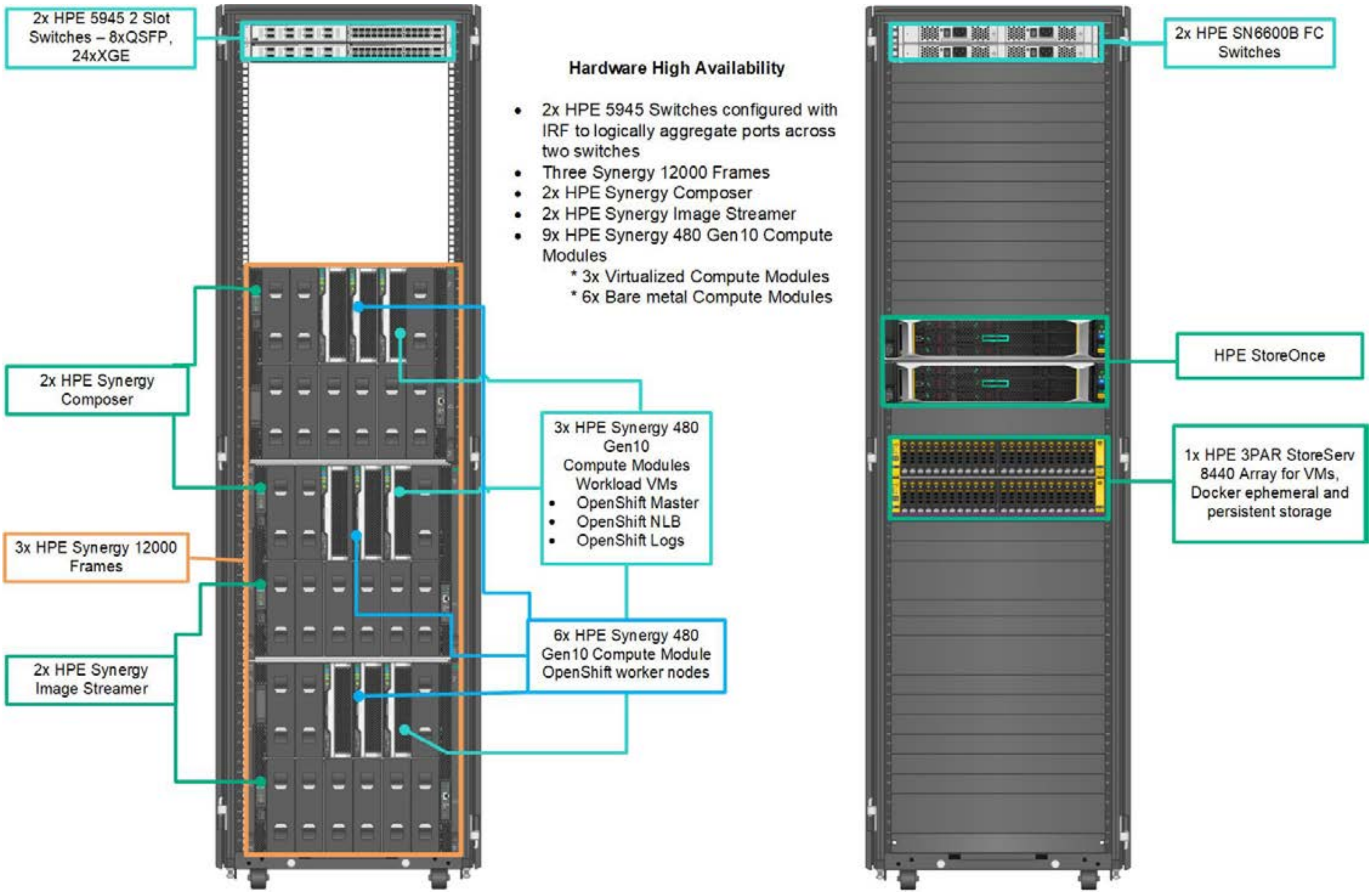


Figure 7. Front view of the solution with individual components highlighted

Solution components

Hardware

The following hardware components were utilized in this Reference Configuration as listed in Table 1.

Table 2. Components utilized in the creation of this solution

Component	Qty	Description
HPE Synergy 12000 Frame	3	Three (3) HPE Synergy 12000 Frames house the infrastructure used for the solution.
• HPE Synergy Composer	2	Two (2) HPE Synergy Composers for core configuration and lifecycle management for the Synergy components.
• HPE Virtual Connect SE 40Gb F8 Module	2	A total of two (2) HPE Virtual Connect SE 40Gb F8 Modules provide network connectivity into and out of the frames.
• HPE Image Streamer	2	Two (2) HPE Image Streamers which provide OS volumes for the OpenShift worker nodes.
HPE Synergy 480 Gen10 Compute Module	9	Three (3) virtualized and six (6) bare metal hosts as described later in this document.
HPE 3PAR StoreServ 8440 Storage	1	One array for virtual machines, Docker storage and persistent volumes.
HPE FlexFabric 5945 2-Slot Switch	2	Each switch contains one (1) each of the HPE 5930 modules listed below:
• HPE 5930 24p SFP+ and 2p QSFP+ Module	2	One (1) module per HPE FlexFabric 2-Slot Switch
• HPE 5930 8p QSFP+ Module	2	One (1) module per HPE FlexFabric 2-Slot Switch
		Fibre Channel connectivity options
HPE Virtual Connect SE 16Gb FC Module	6	Six (6) HPE Virtual Connect SE 16Gb FC Modules (two (2) per frame).
HPE StoreFabric SN6600B 32Gb 48/24 Fibre Channel Switch	2	Fibre channel switches for HPE 3PAR storage.

HPE Synergy

HPE Synergy, the first platform built from the ground up for composable infrastructure, empowers IT to create and deliver new value instantly and continuously. This single infrastructure reduces operational complexity for traditional workloads and increases operational velocity for the new breed of applications and services. Through a single interface, HPE Synergy composes compute, storage and fabric pools into configurations for any application. It also enables a broad range of applications from bare metal to virtual machines to containers, and operational models like hybrid cloud and DevOps. HPE Synergy enables IT to rapidly react to new business demands.

HPE Synergy Frames contain a management appliance called the HPE Synergy Composer which hosts HPE OneView. HPE Synergy Composer manages the composable infrastructure and delivers:

- Fluid pools of resources, where a single infrastructure of compute, storage and fabric boots up ready for workloads and demonstrates self-assimilating capacity.
- Software-defined intelligence, with a single interface that precisely composes logical infrastructures at near-instant speeds; and demonstrates template-driven, frictionless operations.
- Unified API access, which enables simple line-of-code programming of every infrastructure element, easily automates IT operational processes and effortlessly automates applications through infrastructure deployment.

HPE Synergy Composer provides the enterprise-level management to compose and deploy system resources to meet your application needs. This management appliance uses software-defined intelligence to aggregate compute, storage, and fabric resources in a manner that scales to your application needs, instead of being restricted to the fixed ratios of traditional resource offerings. HPE Synergy template-based provisioning enables fast time to service with a single point for defining compute module state, pooled storage, network connectivity, and boot image.

HPE OneView is a comprehensive unifying platform designed from the ground up for converged infrastructure management. A unifying platform increases the productivity of every member of the internal IT team across servers, storage, and networking. By streamlining processes, incorporating best practices, and creating a new holistic way to work, HPE OneView provides organizations with a more efficient way to work. It is designed for open integration with existing tools and processes to extend these efficiencies.

HPE OneView is instrumental for the deployment and management of HPE servers and enclosure networking. It collapses infrastructure management tools into a single resource-oriented architecture that provides direct access to all logical and physical resources of the solution.

Logical resources include server profiles and server profile templates, enclosures and enclosure groups, and logical interconnects and logical interconnect groups. Physical resources include compute modules, interconnects, and storage modules.

HPE OneView offers a uniform way for administrators to interact with resources by providing a RESTful API foundation. The RESTful APIs enable administrators to utilize a growing ecosystem of integrations to further expand the advantages of the integrated resource model that removes the need for the administrator to enter and maintain the same configuration data more than once and keep all versions up to date. It encapsulates and abstracts many underlying tools behind the integrated resource model, so the administrator can operate with new levels of simplicity, speed, and agility to provision, monitor, and maintain the solution.

Within the context of the solution, HPE OneView for Synergy is utilized to:

- Configure the profiles of the HPE Synergy Compute Modules, resulting in a complexity reduction of performing more than 100 manual steps to running one Ansible playbook and a time reduction from an hour to 15 minutes.
- Apply and maintain compliance for firmware across the HPE Synergy infrastructure.
- Configure networking from the HPE Synergy Compute Modules to internal and outbound destinations.

HPE Synergy Image Streamer

HPE Synergy Image Streamer implements rapid image/application changes to multiple compute modules in an automated manner. HPE Synergy Image Streamer works with HPE Synergy Composer to rapidly deploy and update multiple physical compute modules. Operating environment images for bare-metal use might boot directly into a running OS, or VM hosts might perform quick image changeovers. “Infrastructure-as-code” (IaC) capability enables fast delivery of applications and services, including the ability to perform rapid workload switching (using Linux, ESXi, or Microsoft® Windows®). Enhanced profiles provide true stateless images, which integrate the server hardware configuration with operating environment images. Enhanced profiles are stored in redundant image repositories and are automatically integrated for simplicity of use. The unified API enables integration, automation, and customization of operations and applications with HPE Synergy Image Streamer.

HPE Synergy Image Streamer was used in this solution to provide Red Hat Enterprise Linux images to the OpenShift Container Platform worker nodes and to provide VMware vSphere images for virtualized hosts.

HPE Synergy 480 Gen10 Compute Module

The HPE Synergy 480 Gen10 Compute Module delivers an efficient and flexible two-socket server to support the most demanding workloads. Powered by Intel® Xeon® Scalable Family of processors, up to 3TB DDR4, and large storage capacity within a composable architecture. HPE Synergy 480 Gen10 Compute Module:

- Is the most secure server with exclusive HPE Silicon Root of Trust. Protect your applications and assets against downtime associated with hacks and viruses.
- Offers customer choice for greater performance and flexibility with Intel Xeon Scalable Family of processors on the HPE Synergy 480 Gen10 architecture.
- Offers Intelligent System Tuning with processor smoothing and workload matching to improve processor throughput/overall performance up to 8% over previous generation.
- Features a maximum memory footprint of 3TB for large in-memory database and analytic applications.
- Features a hybrid HPE Smart Array for both RAID and HBA zoning in a single controller.

The HPE Synergy 480 Gen10 provides the needed compute to power this solution running both Red Hat Virtualization for the core management pieces of Red Hat OpenShift and Red Hat Enterprise Linux to host the worker nodes.

The bill of materials found in Appendix A of this document outlines the configuration of the HPE Synergy Compute Modules used in this solution.

HPE 3PAR StoreServ 8440 Storage

HPE 3PAR Storage is a scalable Tier-1 all-flash storage with the broadest set of storage features to provide a cloud-ready foundation for enterprises. With architecture built to meet extreme demands of performance, capacity and resiliency, HPE 3PAR delivers over 3M IOPS with sub-milliseconds latency and scaling over 20PB in capacity while also future proofing your investment to take advantage of modern media technologies such as NVMe and Storage Class Memory as they become available. The powerful platform also combines HPE InfoSight predictive analytics capabilities to deliver 99.9999% guaranteed availability.

One of the tenets of the HPE 3PAR StoreServ Storage architecture is its ability to handle unpredictable and mixed workloads which makes it an ideal platform to serve containerized workloads in multi-tenant environments. With features such as Virtual Domains and Priority Optimization Quality of Service (QoS), HPE 3PAR StoreServ Storage can deliver predictable application performance while eliminating any noisy neighbors – a key requirement for consolidation of production workloads in a shared environment.

Now all of the powerful features of HPE 3PAR StoreServ Storage are available for Red Hat OpenShift Container Platform to deliver a variety of use cases for workloads that require persistent storage. The features are exposed to Kubernetes via StorageClass parameters to deliver policy-based storage management for containerized workloads.

Capabilities include:

- Dynamic provisioning of volumes which are by default space efficient, thin provisioned, and are also dedupe and compression capable.
- Acceleration of workloads by enabling SSDs as performance caching tier in a hybrid storage array using adaptive flash cache capability.
- Full copy clones and instantaneous writeable snapshots with auto-expiration and retention abilities for automated lifecycle management of volume copies. DevOps teams can share instant copies of production grade data sets for acceleration of data availability in CI/CD workflows.
- Volumes with Quality of Service (QoS) controls to ensure applications receive the right performance SLA from storage. In shared environment with DevOps, this capability sets guard rails to ensure the production SLAs are unaffected by other workloads.
- Secure multi-tenancy through HPE 3PAR Virtual domains for clear segregation of tenant workloads both at logical and physical levels. Enterprises and service providers can now deliver Container-as-a-Service (CaaS) use cases to multiple customers on the same storage array.
- Support for multiple transport protocols – iSCSI or FC – to suit whatever your environment needs.

HPE StoreOnce

HPE StoreOnce addresses the needs of customers ranging from entry level to large-scale enterprises. HPE StoreOnce systems deliver scale-out capacity and performance to keep pace with shrinking backup windows, reliable disaster recovery, simplified protection of remote offices, and rapid file restore to meet today's SLAs. HPE StoreOnce models vary by capacity and connectivity protocol. It is possible to start with a single HPE StoreOnce base unit or Virtual Storage Appliance (VSA) and then expand with additional units and expansion shelves.

The HPE StoreOnce VSA extends the deployment options for HPE StoreOnce with the agility and flexibility of a virtual appliance, removing the need to install dedicated data protection hardware. All the features of HPE StoreOnce systems are available in a software-defined backup target with up to 500TB of usable capacity. This provides a flexible and cost-effective backup target for virtualized server environments as part of a pure software-defined data protection solution or in conjunction with HPE StoreOnce purpose-built appliances for mixed environments.

HPE StoreOnce Catalyst's data protection optimized interface is unique to HPE StoreOnce Systems. It provides higher performance and more flexible control than traditional emulated tape like Virtual Tape Library (VTL) targets or Network-attached storage (NAS) shares.

Cloud Bank Storage is an extension to StoreOnce Catalyst that combines the low cost object storage with the storage efficiency of HPE StoreOnce deduplication. It connects to external object storage to provide capacity for the Catalyst Cloud Bank store. Using external storage in addition to the local System storage can triple the effective capacity of the StoreOnce System. Through cloud optimized data transfer and storage, Cloud Bank Storage minimizes cloud transfer and storage costs. For more information see, https://h20195.www2.hpe.com/v2/gethtml.aspx?docname=c04328820.

Solution software

Red Hat Enterprise Linux

Red Hat Enterprise Linux Server powers the applications that run your business with the control, confidence, and freedom that comes from a consistent foundation across hybrid cloud deployments. As the premier platform provider for enterprise workloads, Red Hat works side-by-side with engineers from major hardware vendors and cloud providers to make sure that the operating system takes full advantage of the latest innovations. This leadership with partners, as well as Red Hat's influence and contributions to upstream communities, provides a stable, secure, and performance driven foundation for the applications that run the business of today and tomorrow.[2]

This solution is built on Red Hat Enterprise Linux. Each of the Red Hat OpenShift Container Platform control plane nodes, running as virtual machines, are running Red Hat Enterprise Linux as are the dedicated physical servers providing the worker node functions.

VMware vSphere Virtualization

The Red Hat OpenShift Container Platform control plane, including the master, infrastructure, and load balancer roles, are deployed on virtual machines that are distributed across a three-node vSphere cluster. When iSCSI storage connectivity is utilized, it is possible to host worker nodes as virtual machines. When the Storage is FC connectivity, Image registry is hosted on the physical worker nodes

---

Note

Hewlett Packard Enterprise has documented the use of Red Hat Enterprise Virtualization as a hypervisor as part of this solution. Instructions are available at https://github.com/hewlettpackard/hpe-solutions-openshift/tree/master/synergy/scalable/3par-vsphere.

Red Hat OpenShift Container Platform

Red Hat OpenShift Container Platform unites developers and IT operations on a single platform to build, deploy, and manage applications consistently across hybrid cloud and multi-cloud infrastructures. Red Hat OpenShift helps businesses achieve greater value by delivering modern and traditional applications with shorter development cycles and lower operating costs. Red Hat OpenShift is built on open source innovation and industry standards, including Kubernetes and Red Hat Enterprise Linux, the world's leading enterprise Linux distribution.[3]

HPE Recovery Manager Central Software

HPE Recovery Manager Central (RMC) software facilitates policy-driven, converged data protection and copy data management for business-critical applications at speeds required for all-flash storage. RMC integrates HPE 3PAR StoreServ and HPE Nimble All-Flash arrays with HPE StoreOnce Systems, leveraging snapshot performance with storage-integrated backups to deliver flash speed application protection and copy data management with less cost and complexity than legacy solutions. RMC is also built for cloud, allowing users to leverage public cloud for cost-effective, long-term retention of user's backups. For more information visit https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-c04552659.

Sysdig Agent

The Sysdig Agent is installed as part of the Sysdig cloud-native intelligence platform, Sysdig Secure and Sysdig Monitor, in order to offer unified container security, monitoring, and forensics for container and Kubernetes based environments. Sysdig Agent moves the data collected to the Sysdig SaaS Platform where the security and monitoring services will be available to the user based on the subscription model they have selected.

[2] redhat.com/cms/managed-files/li-enterprise-linux-server-datasheet-f11191jm-201803-en.pdf
[3] redhat.com/cms/managed-files/cl-OpenShift-container-platform-datasheet-f9695kc-201711-en.pdf

Red Hat OpenShift Container Platform layout

Figure 8 highlights how the individual Red Hat OpenShift Container Platform pieces are laid out within the solution. Optionally, when iSCSI storage connectivity is used, virtual worker nodes can also be deployed as part of the solution.

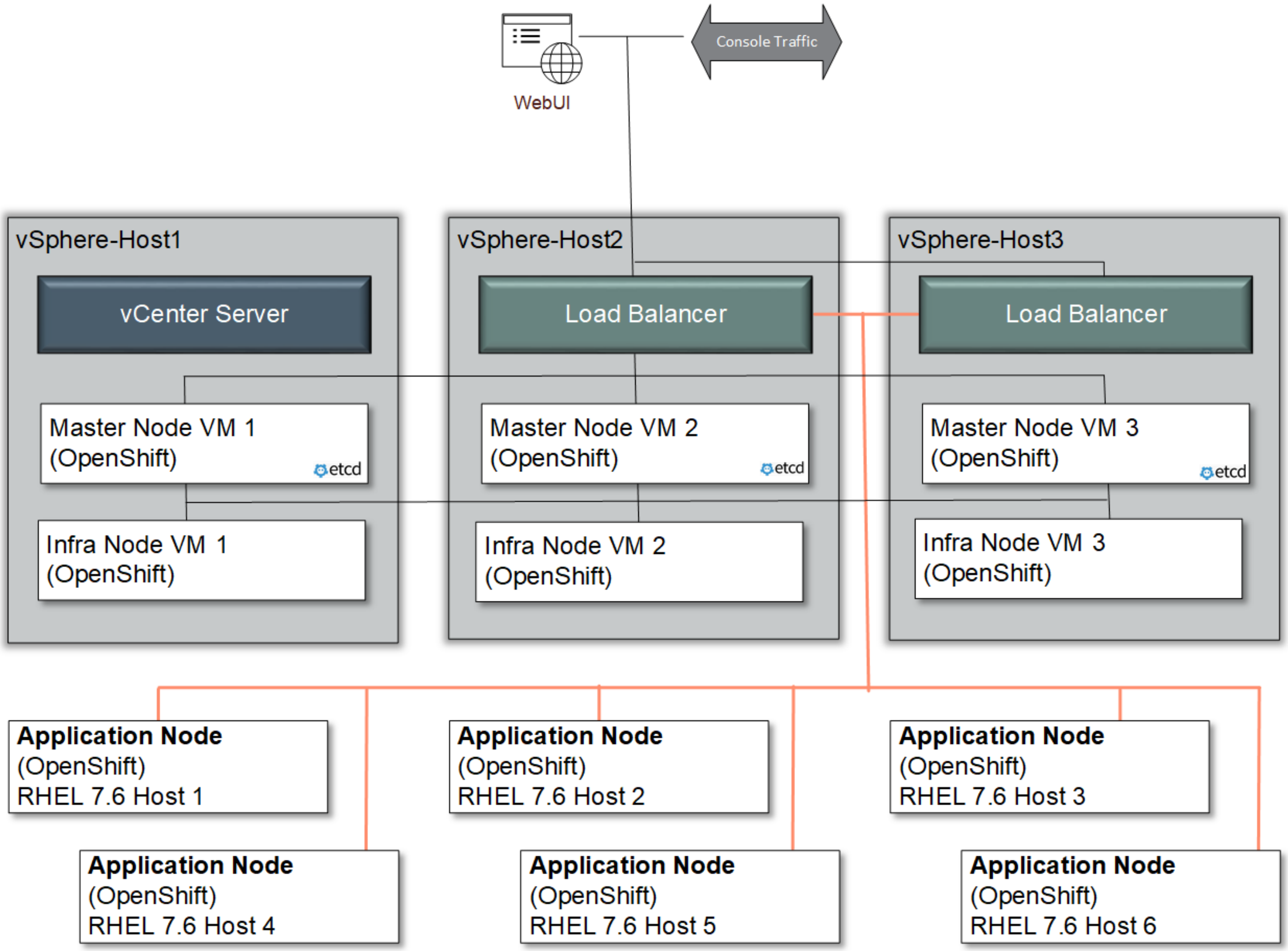


Figure 8. Red Hat OpenShift Container Platform layout

Best practices and configuration guidance for the solution

This section discusses the high-level cabling and configuration of the solution hardware and software. For a detailed explanation of how to build and deploy the entire solution stack, consult the deployment guide and accompanying Ansible deployment playbooks at https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/3par-vsphere.

Solution cabling

Figure 9 below describes the cabling configuration of the three (3) Synergy 12000 Frames as well as the HPE FlexFabric 5945 switches and Intelligent Resilient Fabric (IRF) within the context of this solution. These cables carry frame management, inter-frame and interconnect traffic between frames.

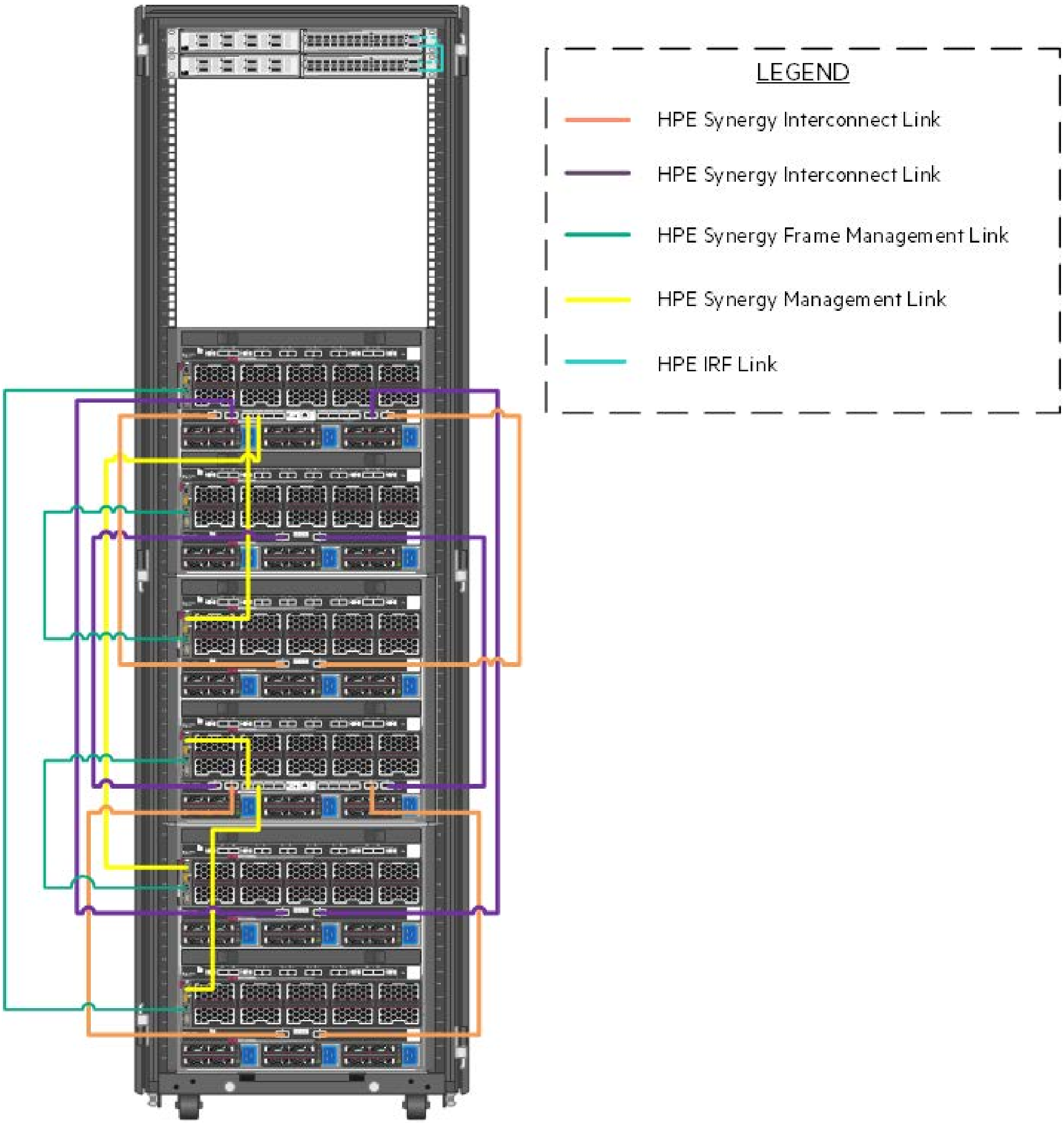


Figure 9. Frame and switch cabling within the solution

Networking

Figure 10 documents the cabling of the solution from the HPE Virtual Connect SE 40 GB modules to the switches as well as the traffic carried on each connection. While top of rack switching was used in the creation of this solution, end of row switching is generally preferred in HPE Synergy environments. This results in a reduction in the number of physical switches as the solution scales to support more HPE Synergy Frames. Top of rack was used for solution development to facilitate the connectivity to the HPE Storage iSCSI NICs in an adjacent rack. Switching is flexible based on the installation environment. In case of Fibre Channel storage only, data center networks will be configured. Data traffic will be over FC.

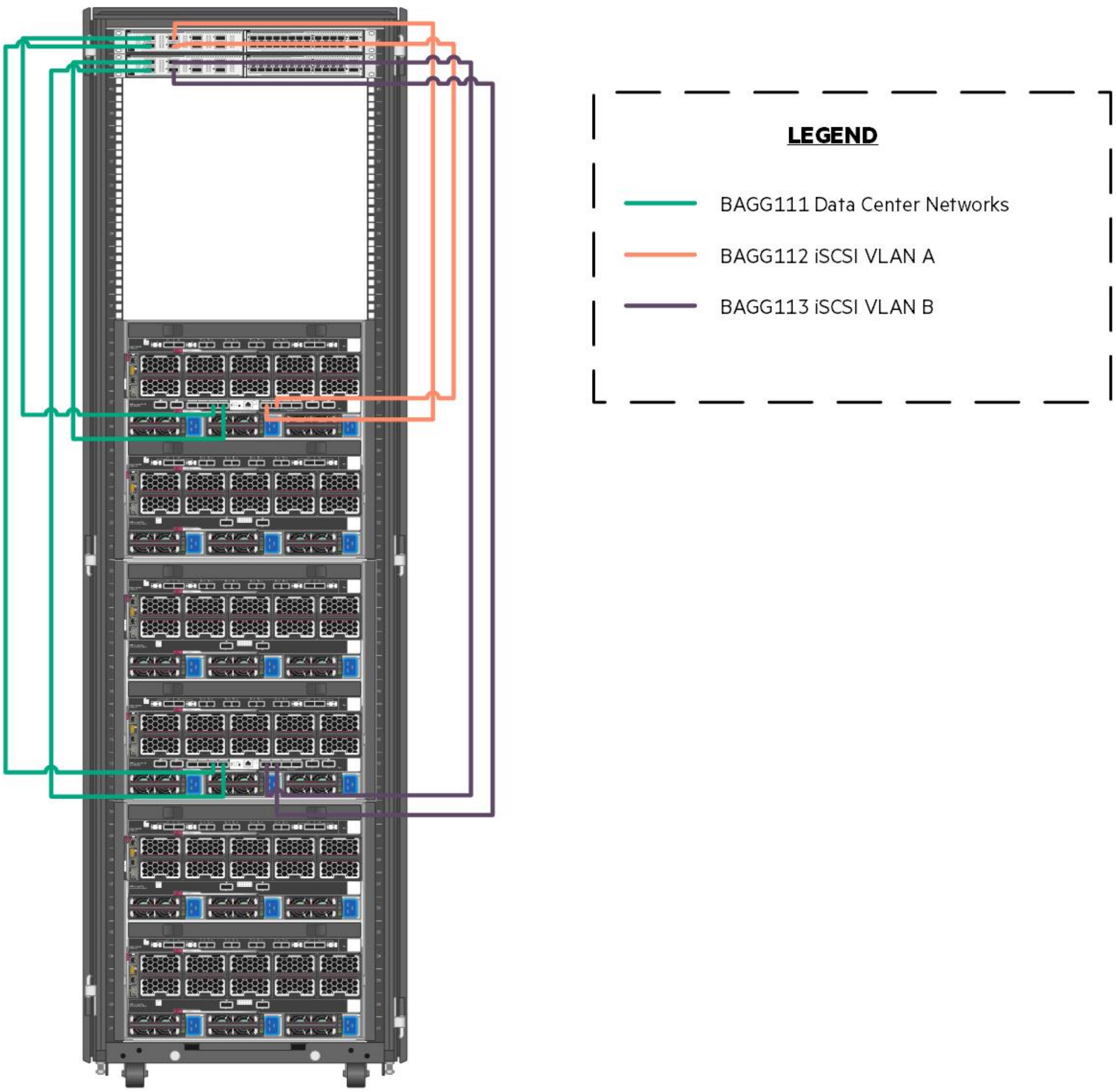


Figure 10. Network cabling from the HPE Synergy 12000 Frames to the switches

Table 3 describes the configuration of the networks as defined within HPE OneView for Synergy and the bandwidth associated with each network. iSCSI VLAN's are configured in case of iSCSI storage deployment only.

Table 3. Networks defined within HPE Synergy Composer for this solution

Network Name	Type	VLAN Number	Purpose	Requested Bandwidth (Gb)	Maximum Bandwidth (Gb)
Management	Ethernet	1193	Solution management	5	20
Data_Center	Ethernet	2193	Application access, authentication and other user networks	5	20
Synergy Management	Ethernet	193	Image Streamer, OneView for Synergy	1	20
ISCSI_VLAN_A	Ethernet	3193	ISCSI VLAN A	8	20
ISCSI_VLAN_B	Ethernet	3194	ISCSI VLAN B	8	20
Deployment	Ethernet	500	Deployment Network	2	20

Table 4 describes the cabling of the HPE Virtual Connect SE 40Gb F8 Modules for Synergy to the switches and highlights what networks are carried on the connections. All Ethernet networks as described in Table 3 are carried on the Uplink Set labeled “Network”. iSCSI is configured in case of iSCSI storage deployment only.

Table 4. Networks used in this solution

Uplink Set	Synergy Source	Switch Destination
Network	Enclosure 1 Port Q3	FortyGigE1/1/1
	Enclosure 1 Port Q4	FortyGigE2/1/1
	Enclosure 2 Port Q3	FortyGigE1/1/2
	Enclosure 2 Port Q4	FortyGigE2/1/2
iSCSI_SAN_A	Enclosure 1 Port Q5	FortyGigE1/1/5
	Enclosure 1 Port Q6	FortyGigE1/1/6
iSCSI_SAN_B	Enclosure 2 Port Q5	FortyGigE2/1/5
	Enclosure 2 Port Q6	FortyGigE2/1/6

By utilizing HPE Synergy, the non-storage networks within the solution are able to traverse the Synergy infrastructure in an east-west fashion across high speed, low latency links both within and between HPE Virtual Connect Modules. In particular, communication between the OpenShift management pieces remains within the HPE Synergy Frames.

By utilizing HPE Synergy, the non-storage networks within the solution are able to traverse the Synergy infrastructure in an east-west fashion across high speed, low latency links, both within and between HPE Virtual Connect Modules. In particular, communication between the core OpenShift management pieces remains within the HPE Synergy Frames. Figure 11 and 12 shows the SAN configuration in context of the HPE Synergy Compute Modules utilized for management functions for both FC and iSCSI. In this part of the solution, the compute modules are virtualized and as such, Ethernet traffic traverses a virtual switch prior to the Synergy Link Aggregation Group (LAG).

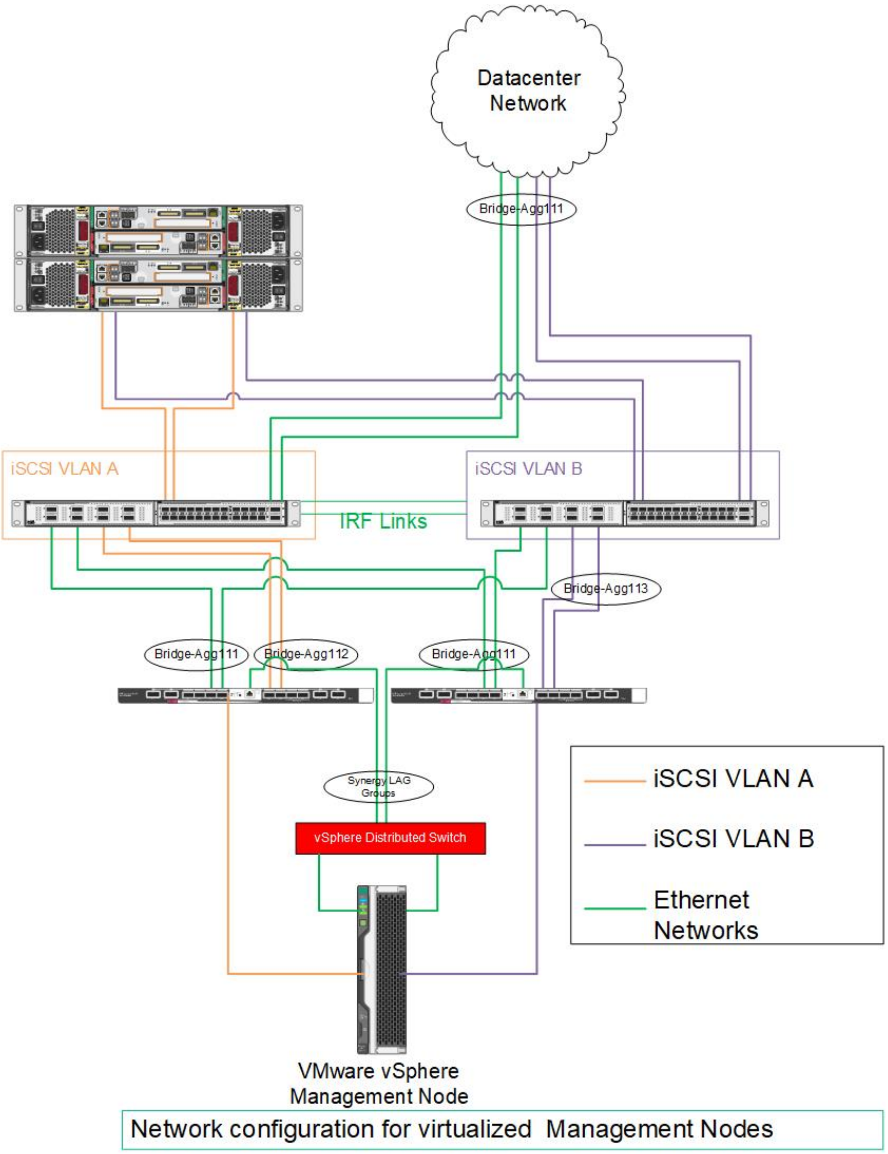


Figure 11 iSCSI configuration for virtualized management hosts -iSCSI

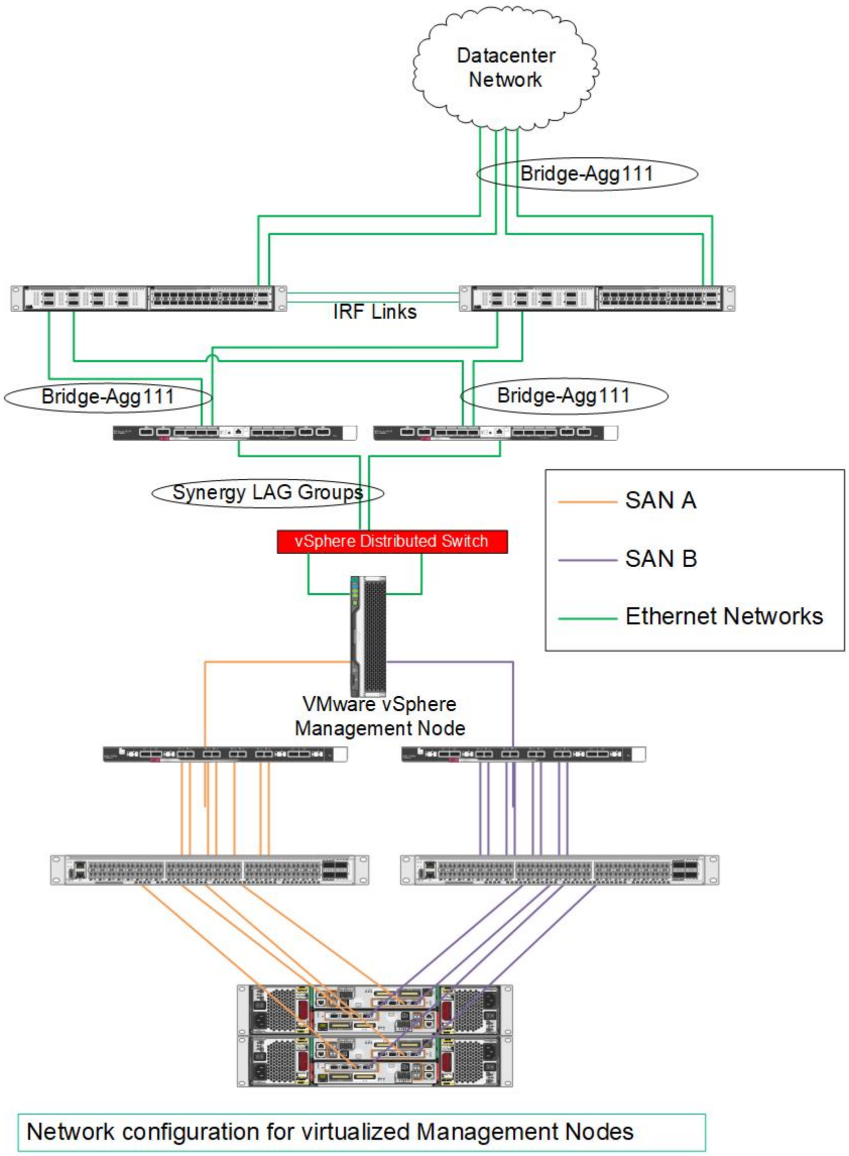


Figure 12 SAN configuration for virtualized management hosts -FC

Figure 13 represents the SAN diagram in the context of the OpenShift worker nodes for iSCSI. These compute modules run Red Hat Enterprise Linux. A single storage array is connected in this diagram. iSCSI traffic does not pass through a LAG, but rather is untagged from source to final destination.

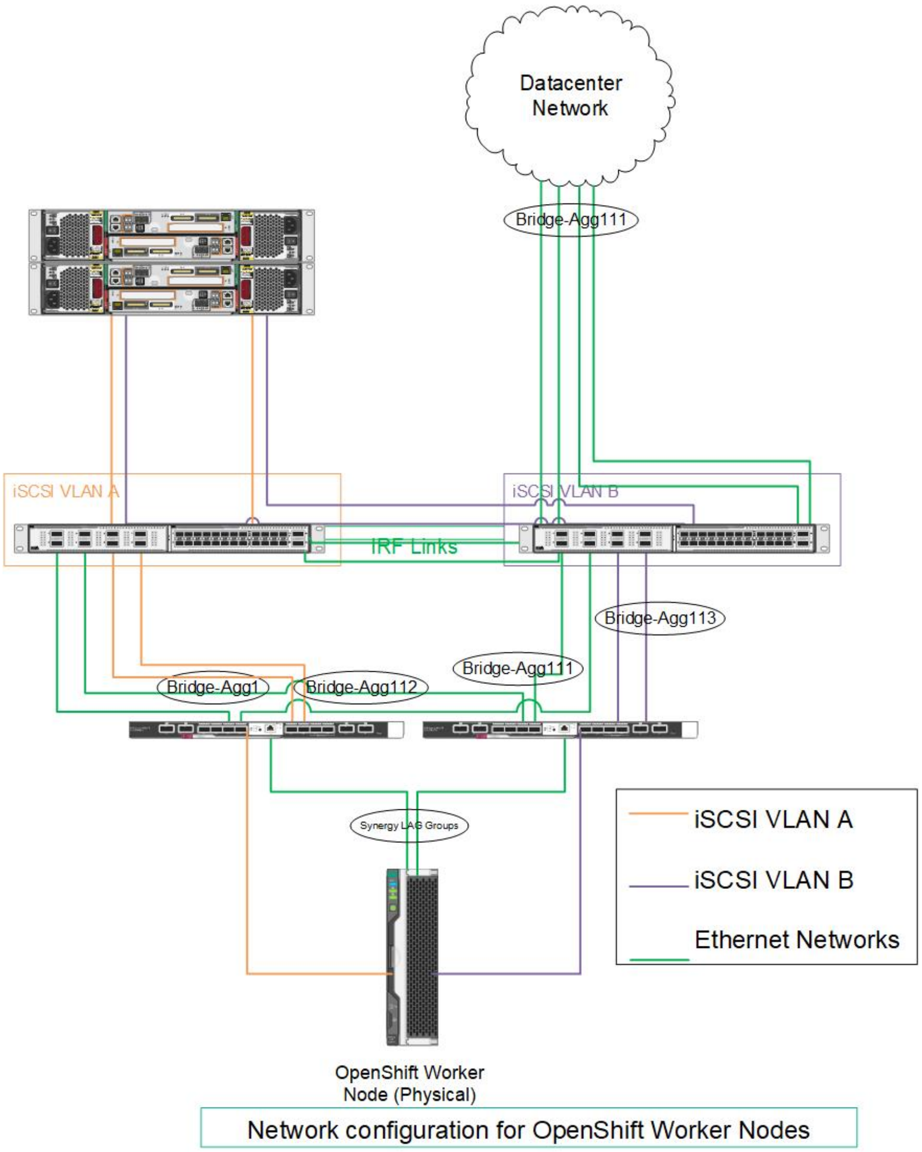


Figure13. SAN configuration for worker nodes connected via iSCSI

Figure 14 presents the same view of SAN connectivity when systems are connected via Fibre Channel rather than iSCSI. This requires the addition of Fibre Channel switching as well as HPE Virtual Connect Fibre Channel modules.

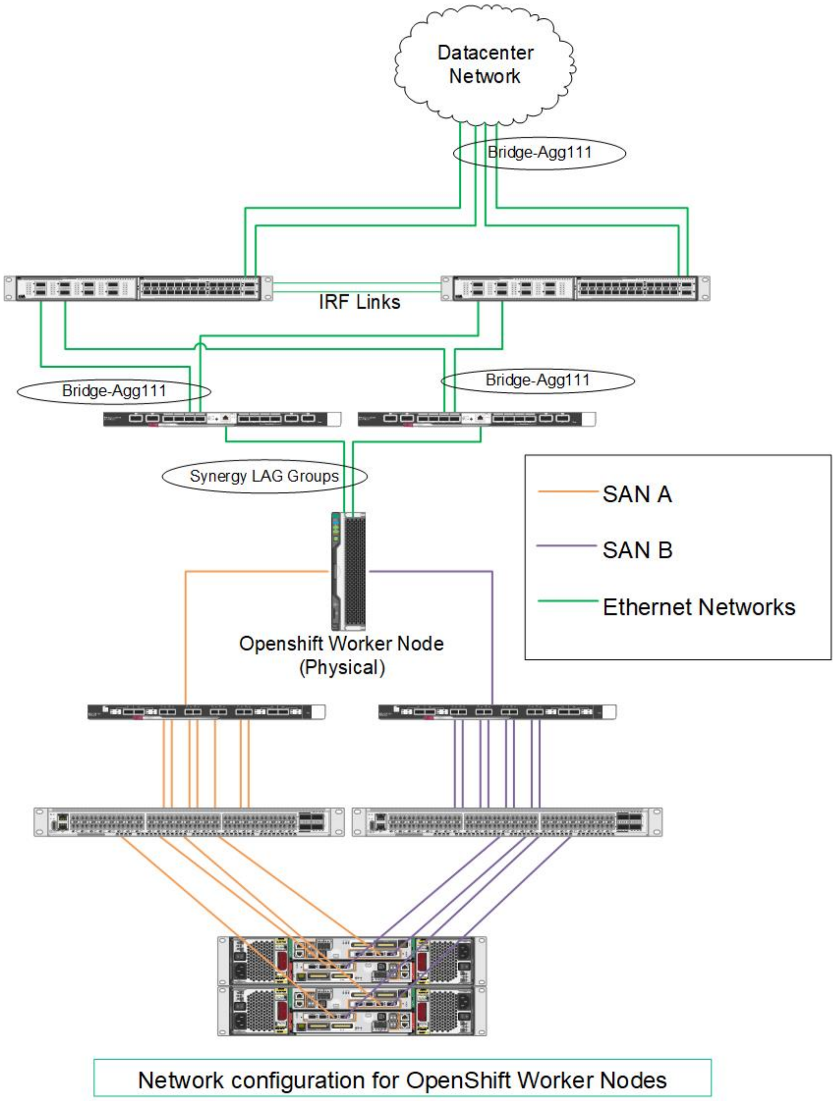


Figure 14. SAN configuration for worker nodes connected via Fibre Channel

Storage

This section describes the configuration of the HPE 3PAR StoreServ 8440 storage in the context of this solution. The HPE 3PAR StoreServ 8440 provides both shared and dedicated storage for:

- Docker local storage
- Virtual Machines including storage for VMWare vSphere
- Persistent volumes

Figure 15 and 16 describes the logical storage layout of iSCSI and FC respectively, used in the solution. Local storage is used for the operating system installation on Red Hat Enterprise worker hosts while HPE Image Streamer provides the OS volumes to the virtualized VMware vSphere management nodes, and the bare metal worker nodes. In the case of iSCSI storage deployment, HPE 3PAR StoreServ 8440 provides dedicated and shared volumes as outlined in the Figure 15.

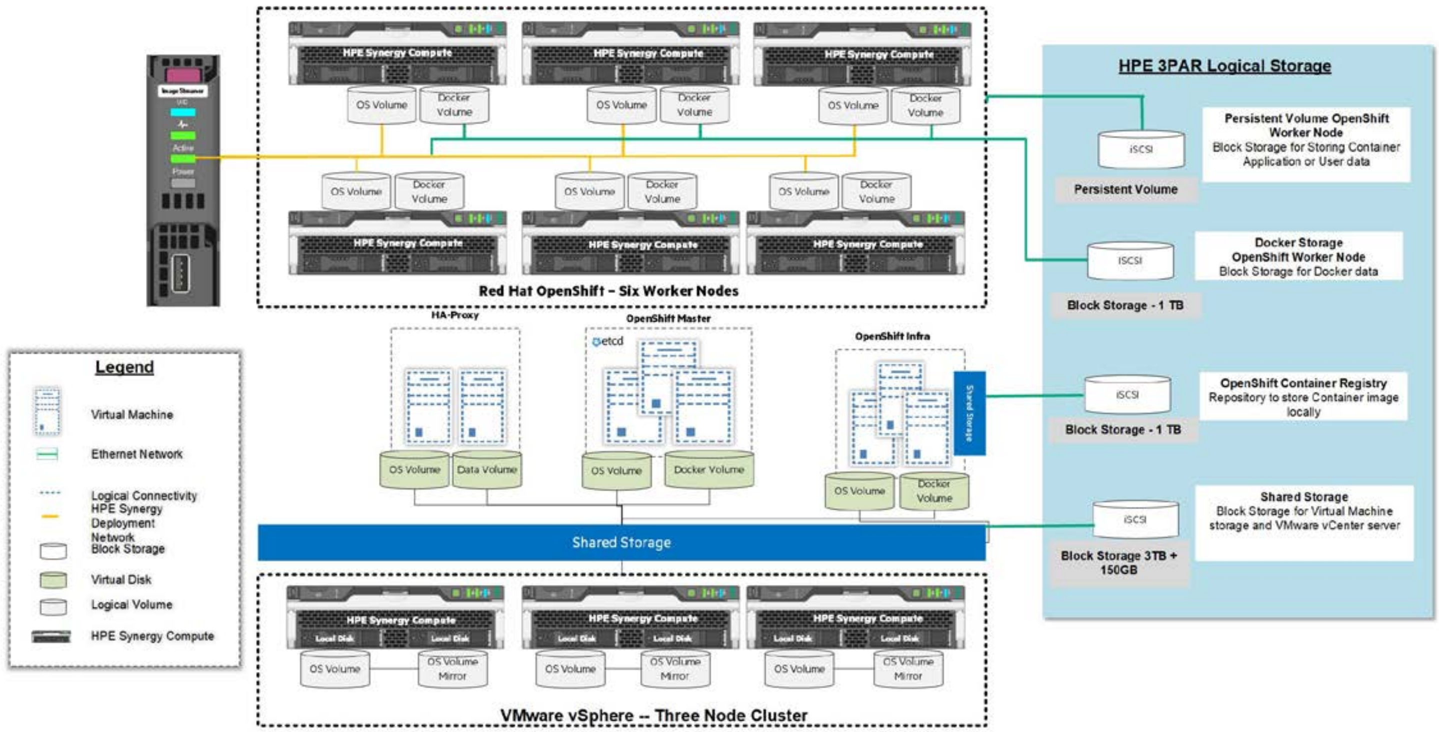


Figure 15. Logical storage layout within the solution -iSCSI

In the case of FC storage deployment, HPE 3PAR StoreServ 8440 provides dedicated and shared volumes as outlined in the Figure16.

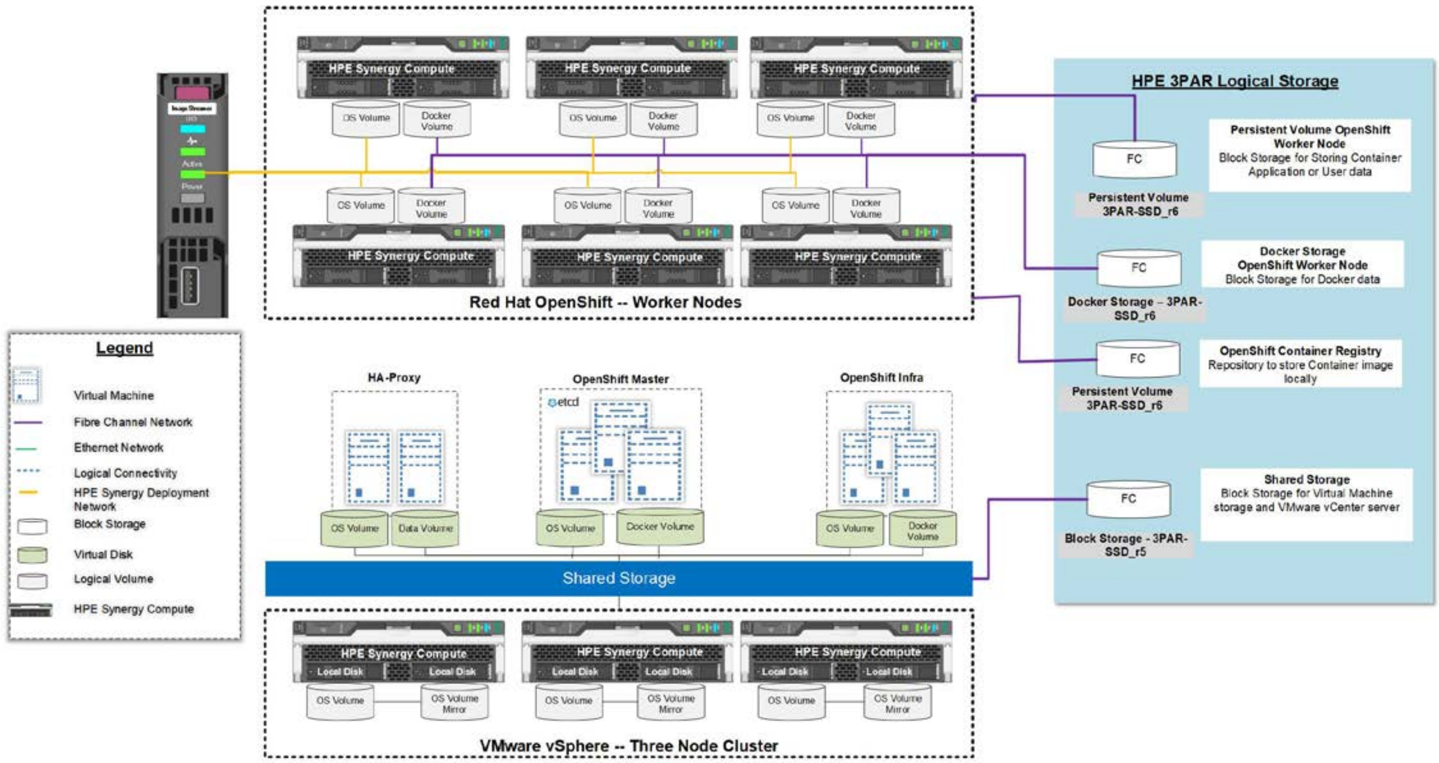


Figure 16. Logical storage layout within the solution -FC

Table 5 lists all volumes used within the solution and highlights what storage provides the capacity and performance for each function. This is applicable for both iSCSI and FC storages.

Table 5. Volumes used in this solution

Volume/Disk Function	Qty	Size	Source	Hosts	Shared/Dedicated
Hypervisor	3	40GB	HPE Image Streamer	VMWare vSphere hosts	Dedicated
Operating System	6	40GB	HPE Image Streamer	OpenShift worker nodes	Both
Virtual Machine Hosting	1	3TB	HPE 3PAR StoreServ	VMware vSphere hosts	Shared
Persistent Application Data	N	App Specific	HPE 3PAR StoreServ	Worker nodes	Dedicated
OpenShift Container registry	1	1TB	HPE 3PAR StoreServ	Infrastructure nodes Worker nodes (for FC storage)	Shared
Docker Local Storage	6	1TB	HPE 3PAR StoreServ	Worker nodes	Dedicated
VMware vCenter Server	1	150GB	HPE 3PAR StoreServ	vSphere hosts or external	Shared

Note

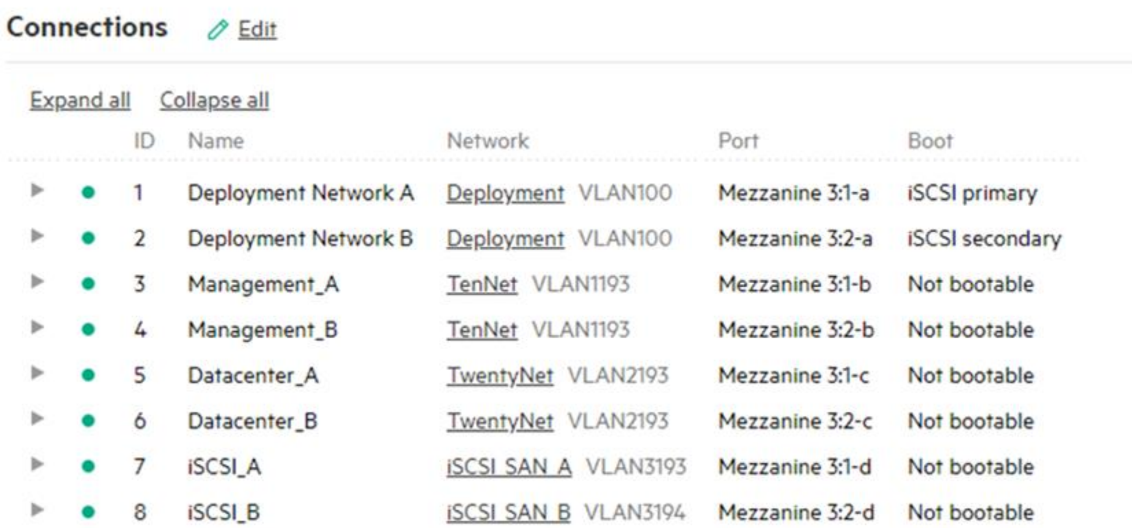
In case of FC storage OpenShift Container registry role is assigned with worker nodes and the external storage is shared among worker nodes.

The array used in the creation of this solution was built to suit configuration testing. Customer requirements around application performance and capacity should be taken into account when selecting an HPE 3PAR StoreServ array for production Red Hat OpenShift Container Platform environments.

Server profiles

HPE Synergy Composable Infrastructure using HPE Virtual Connect provides the construct of a server profile. A server profile allows a suite of configuration parameters, including network and SAN connectivity, BIOS tuning, boot order configuration, local storage configuration and more to be templatized and applied programmatically to compute resources. These templates are the key to delivering the "infrastructure-as-code" capabilities of the HPE Synergy platform. For the purpose of this solution, a single template was created that was applied to OpenShift worker compute modules.

The critical items configured as part of a template supporting Red Hat OpenShift Container Platform are the network connections and storage. Figure 17 describes the configuration of the network interfaces as part of the profile template for the virtualized management as well as the physical worker nodes. There are 8 redundant Ethernet networks that are defined as part of iSCSI Storage. Figure 18 describes the configuration of the network interfaces as part of the profile template for the virtualized management as well as the physical worker nodes. There are 6 redundant Ethernet networks and 2 Fibre Channel connections that are defined as part of FC Storage.


Connections Edit

Expand all Collapse all

ID	Name	Network	Port	Boot
1	Deployment Network A	Deployment VLAN100	Mezzanine 3:1-a	iSCSI primary
2	Deployment Network B	Deployment VLAN100	Mezzanine 3:2-a	iSCSI secondary
3	Management_A	TenNet VLAN1193	Mezzanine 3:1-b	Not bootable
4	Management_B	TenNet VLAN1193	Mezzanine 3:2-b	Not bootable
5	Datacenter_A	TwentyNet VLAN2193	Mezzanine 3:1-c	Not bootable
6	Datacenter_B	TwentyNet VLAN2193	Mezzanine 3:2-c	Not bootable
7	iSCSI_A	iSCSI SAN A VLAN3193	Mezzanine 3:1-d	Not bootable
8	iSCSI_B	iSCSI SAN B VLAN3194	Mezzanine 3:2-d	Not bootable

Figure 17. Server connections as part of the profile for the worker nodes

Connections Edit

Expand all Collapse all

		ID	Name	Network	Port	Boot
▶	●	1	Deployment01	Deployment VLAN100	Mezzanine 3:1-a	iSCSI primary
▶	●	2	Deployment02	Deployment VLAN100	Mezzanine 3:2-a	iSCSI secondary
▶	●	3	Management01	TenNet VLAN1193	Mezzanine 3:1-c	Not bootable
▶	●	4	Management02	TenNet VLAN1193	Mezzanine 3:2-c	Not bootable
▶	●	5	DCNet01	TwentyNet VLAN2193	Mezzanine 3:1-d	Not bootable
▶	●	6	DCnet02	TwentyNet VLAN2193	Mezzanine 3:2-d	Not bootable
▶	●	7	SAN_A	Fiber Channel A Fabric attach	Mezzanine 2:1	Not bootable
▶	●	8	SAN_B	Fiber Channel B Fabric attach	Mezzanine 2:2	Not bootable

Figure 18. Server connections as part of the profile for the worker nodes in case of FC storage

Figure 19 describes the SAN connections of an individual profile for iSCSI. Each server has a minimum of 8Gb/s of bandwidth available for iSCSI adapters.

▼ ●	7	iSCSI_A	iSCSI SAN A VLAN3193	Mezzanine 3:1-d	Not bootable
	Interconnect		2S1721PK4K, interconnect 3		
	Type		Ethernet		
	MAC address		06:FB:29:B0:0D:56 (v)		
	Requested virtual functions		None		
	Requested bandwidth		8 Gb/s		
	Allocated bandwidth		8 Gb/s		
	Max bandwidth		20 Gb/s		
	Link aggregation group		*None*		
▼ ●	8	iSCSI_B	iSCSI SAN B VLAN3194	Mezzanine 3:2-d	Not bootable
	Interconnect		MXQ73007JR, interconnect 6		
	Type		Ethernet		
	MAC address		06:FB:29:B0:0D:57 (v)		
	Requested virtual functions		None		
	Requested bandwidth		8 Gb/s		
	Allocated bandwidth		8 Gb/s		
	Max bandwidth		20 Gb/s		
	Link aggregation group		*None*		

Figure 19. iSCSI SAN connections as viewed from the server profile.

Figure 20 represents the server profile configuration for SAN on Fibre Channel connected hosts. There is 16Gb/s of bandwidth available per FC HBA.

7	SAN_A	Fiber_Channel_A Fabric attach	Mezzanine 2:1	Not bootable
	Interconnect	2S1721PK4K, interconnect 2		
	Type	Fibre Channel		
	WWNN	10:00:56:a9:fa:80:00:f7 (v)		
	WWPN	10:00:56:a9:fa:80:00:f6 (v)		
	MAC address	06:FB:29:B0:04:28 (v)		
	Requested bandwidth	Auto		
	Allocated bandwidth	16 Gb/s		
	Max bandwidth	16 Gb/s		
8	SAN_B	Fiber_Channel_B Fabric attach	Mezzanine 2:2	Not bootable
	Interconnect	2S1721PK4K, interconnect 5		
	Type	Fibre Channel		
	WWNN	10:00:56:a9:fa:80:00:f9 (v)		
	WWPN	10:00:56:a9:fa:80:00:f8 (v)		
	MAC address	06:FB:29:B0:04:29 (v)		
	Requested bandwidth	Auto		
	Allocated bandwidth	16 Gb/s		
	Max bandwidth	16 Gb/s		

Figure 20. SAN connections as viewed from the server profile.

Red Hat OpenShift virtual machines

Red Hat OpenShift Container Platform requires a number of redundant functions. These functions may be hosted on either physical compute modules or on virtual machines, both of which run Red Hat Enterprise Linux version 7.6. For this solution, Hewlett Packard Enterprise choose to implement these functions as virtual machines. This approach reduces the amount of infrastructure required while introducing enhanced options for management and high availability. Three (3) HPE Synergy 480 Gen10 Compute Modules host the virtual machines as shown in Figure 21 and 22 below. These figures also show that the worker nodes run on bare metal. While the solution was tested with six (6) worker nodes, it is scalable to include many more. While worker nodes are depicted as physical resources in Figure 21 and 22, the hosts can be deployed as hypervisors and the workers can be deployed as virtual machines in variable quantities according to the container requirements.

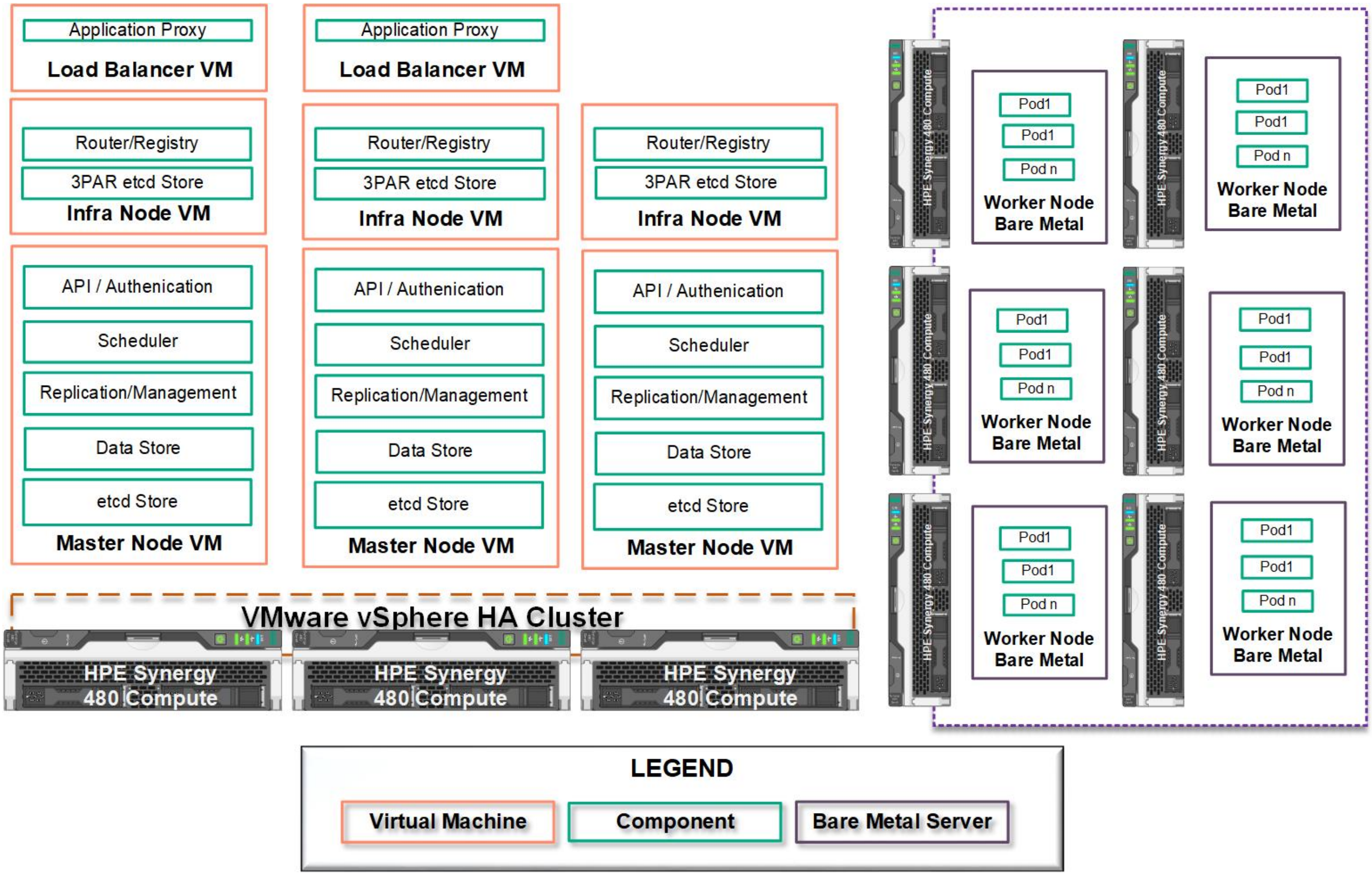


Figure 21. Red Hat OpenShift implementation with virtual machines for core management functions and iSCSI storage connectivity

For the FC solution, persistent storage volumes were targeted at OpenShift nodes deployed on physical servers rather than virtual machines. To provide persistent storage backing for the OpenShift registry, the region=infra selector label was applied to the worker nodes running on physical servers. This ensures the registry is deployed on a physical server that is capable of providing persistent storage. This in turn eliminates the need for dedicated infrastructure VMs running within the solution.

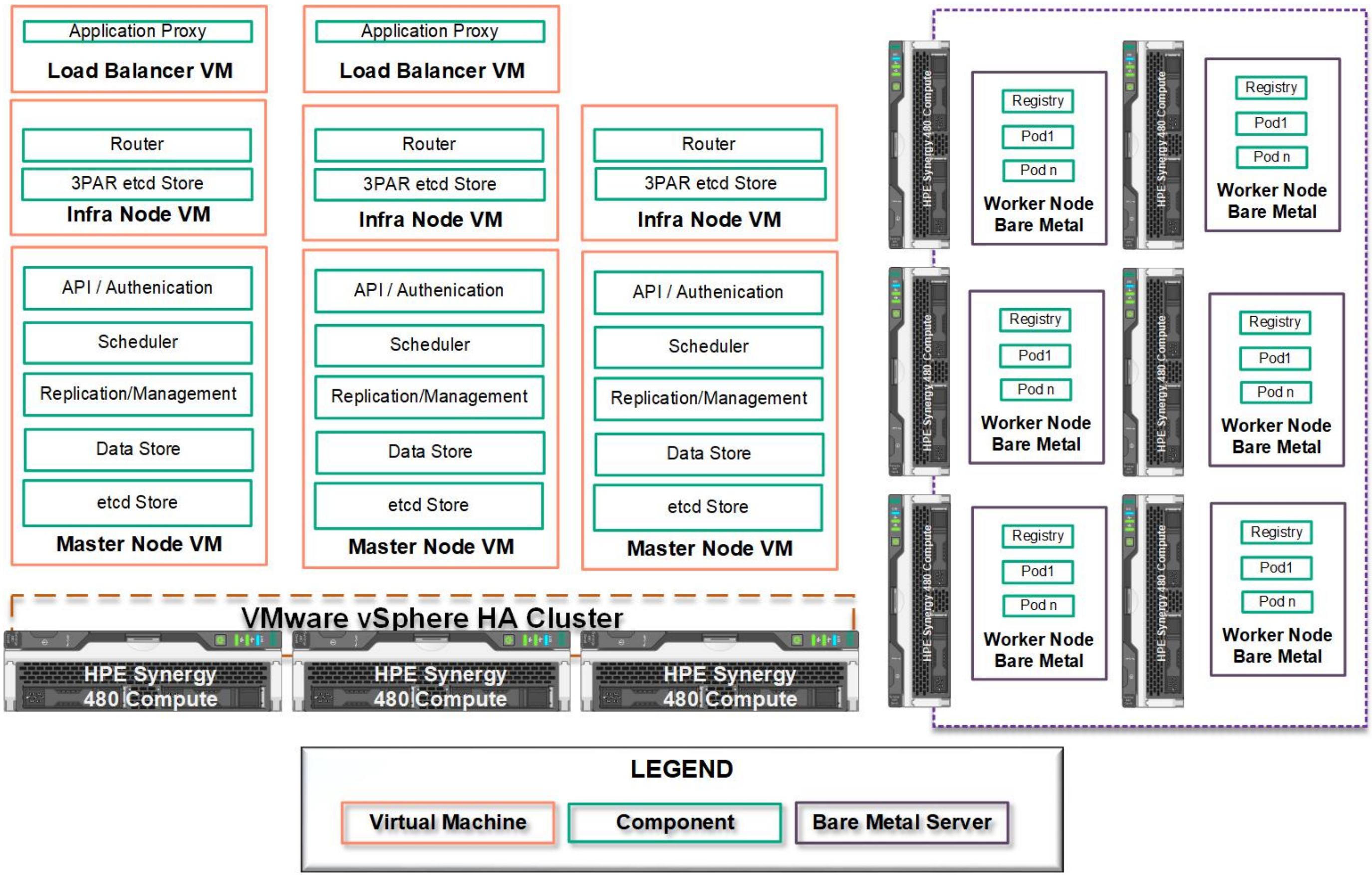


Figure 22. Red Hat OpenShift implementation with virtual machines for core management functions and Fibre Channel storage connectivity

The virtual environment is managed by VMware vCenter Server which runs as a virtual machine within the environment. This virtual machine is deployed to a dedicated volume hosted on HPE 3PAR StoreServ Storage.

Software

This section describes the software versions utilized in the solution as well as noting any special installation or configuration requirements. For detailed descriptions on how to install and configure the software, refer to the deployment guide and accompanying scripts found at, https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/3par-vsphere.

Table 6 lists the versions of required software used in the creation of this solution.

Table 6. Software versions used in the solution

Component	Version
Red Hat Enterprise Linux Server	7.6
vCenter, vSphere	6.7
Red Hat OpenShift Container Platform	3.11
Python	2.7.9 or above
Ansible version	2.7.2
Sysdig Agent	0.90.2

Capacity and sizing

Sizing for a Red Hat OpenShift Container Platform environment varies depending on the requirements of the specific organization and type of deployment. In this section, we will discuss sizing considerations for Red Hat OpenShift Container Platform virtual machines, host requirements, and cluster sizing.

Red Hat OpenShift Container Platform role sizing

- Master – The minimum size for a physical or virtual machine running the master node is 4 vCPU and 16 GB RAM with a 50 GB disk space for /var, 1 GB disk space for /usr/local/bin, and 1 GB disk space for the system's temporary directory. Master nodes should be configured with an additional 1 CPU core and 1.5 GB RAM for each additional 1000 pods. There are three (3) total master node virtual machines in this solution.
- HAProxy – Two (2) HAProxy load balancer virtual machines were deployed in this solution.
- Ansible Engine + Web server – The minimum size for the Ansible Engine & web server virtual machine node is 4 vCPU and 250 GB RAM. The minimum partition size for a web server virtual machine is 150 GB RAM for "/" directory for downloading required software into webserver and 100 GB is sufficient for managing Ansible Engine.

Red Hat OpenShift Container Platform cluster sizing

The number of application nodes in an OpenShift Cluster depends on the number of pods that an organization is planning on deploying. Red Hat OpenShift Container Platform version 3.11 can support the following maximums.

- Maximum of 2000 nodes per cluster
- Maximum of 150,000 pods per cluster
- Maximum of 250 pods per node
- Maximum of 10 pods per CPU core

To determine the number of nodes required in a cluster, estimate the number of pods the organization is planning on deploying and divide by the maximum number of pods per node. For example, if the organization expects to deploy 5000 pods, then the organization should expect to deploy 20 application nodes with 250 pods per node (5000 / 250 =20). In this environment with a default configuration of three physical application nodes, the Red Hat OpenShift Cluster should be expected to support 750 pods (250 pods x 3 nodes = 750 pods).

For more information about Red Hat OpenShift Container Platform sizing, refer to the Red Hat OpenShift Container Platform product documentation at, https://docs.OpenShift.com/container-platform/3.11/scaling_performance/index.html.

Solution deployment overview

In depth details of solution deployment are documented in the deployment guide and accompanying Ansible automation scripts at https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/nimble. This document also outlines about configuring OpenShift in a disconnected environment, Configuring Sysdig security, and Configuring Data protection for Red Hat OpenShift Cluster.

Figure 23 shows the overall deployment flow. For more detailed flow instructions, refer to the deployment guide at https://github.com/HewlettPackard/hpe-solutions-OpenShift/tree/master/synergy/scalable/3par-vsphere.

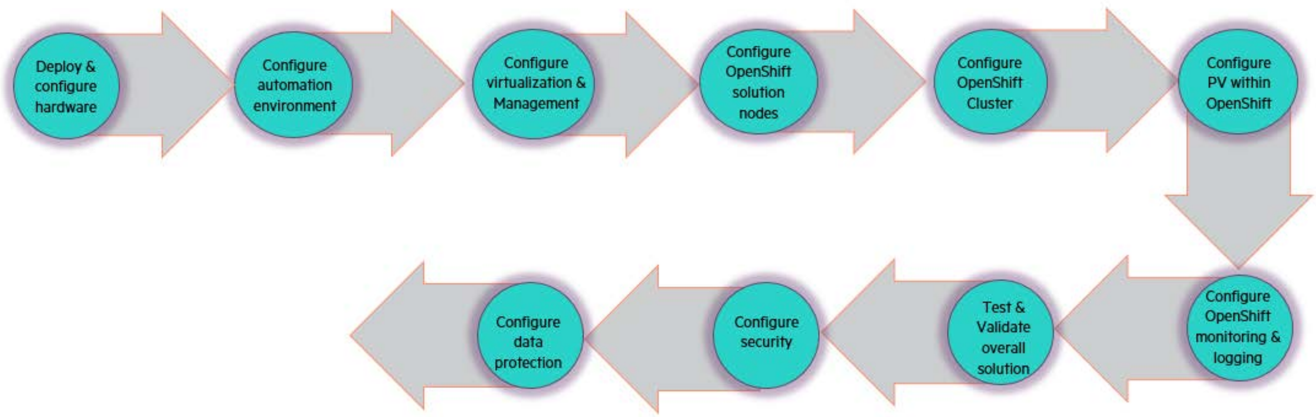


Figure 23. Red Hat OpenShift solution setup overview

Summary

Red Hat OpenShift Container Platform on HPE Synergy provides an end-to-end, fully integrated container solution that, once assembled, can be configured within hours. This eliminates the complexities associated with implementing a container platform across an enterprise data center and provides the automation of hardware and software configuration to quickly provision and deploy a containerized environment at scale. Red Hat OpenShift Container Platform provides organizations with a reliable platform for deploying and scaling container-based applications and HPE Synergy provides the flexible infrastructure you need to run that container platform to dynamically provision and scale applications, whether they run in VMs or containers, or are hosted on-premises in the cloud, or somewhere in between.

This Reference Configuration demonstrated the following benefits of utilizing HPE Synergy for Red Hat OpenShift Container Platform:

- Automated initial installation and configuration of highly available vSphere hosts and the management virtual machine for the Red Hat OpenShift Container Platform, is reduced from more than 3 hours to approximately one hour, and complexity of the manual operation is reduced from needing to perform more than 500 steps, to running two Ansible play books.
- Automated deployment and configuration of physical worker nodes is reduced from 8 hours to under 20 minutes, and complexity of manual operation is reduced from needing to perform close to 300 steps, to running two Ansible play books.
- Automated host preparation of the Red Hat OpenShift Container Platform Nodes is reduced from up to 8 hours to as little as 20 minutes, and the complexity of the manual operation is reduced from performing more than three hundred steps, to one Ansible playbook.
- Deploying the management and load balancers on VMs, to optimize resource usage, while keeping the worker nodes on bare metal to optimize for performance.
- Using an enterprise grade storage solution such as HPE 3PAR StoreServ Storage for Persistent Storage with containers that enables speed, portability, and agility for traditional enterprise applications and data.
- The HPE Composable Infrastructure solution provides a layered view of security controls. The objective of choosing this layered security view is to ensure that the customers are aware of the depth of security risk that an infrastructure can have and also make them aware of the depth of defense that is built-in the HPE Composable Infrastructure design.
- Container Platform security provided by Sysdig cloud-native intelligence platform, Sysdig Secure and Sysdig Monitor, offer unified container security, monitoring, and forensics for container and Kubernetes based environments.
- Data protection for container persistent data is provided by HPE RMC and HPE StoreOnce.
- Provides deployment capabilities for Red Hat OpenShift Container Platform in an offline fashion.

Appendix A: Bill of materials

The following bill of materials contains the core components utilized in the creation of this solution. Services, support and software are not included in the BOM and, along with power distribution, should be customized based on customer needs.

Note

Part numbers are at time of testing and subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your HPE Reseller or HPE Sales Representative for more details. hpe.com/us/en/services/consulting.html

---

Table A1. Bill of materials

Qty	Part number	Description
		Rack and Network Infrastructure
1	P9K10A	HPE 42U 600mmx1200mm G2 Kitted Advanced Shock Rack with Side Panels and Baying
1	P9K10A 001	HPE Factory Express Base Racking Service
1	H6J85A	HPE Rack Hardware Kit
1	BW932A	HPE 600mm Rack Stabilizer Kit
1	BW932A B01	HPE 600mm Rack include with Complete System Stabilizer Kit
4	AF533A	HPE Intelligent Modular 3Ph 14.4kVA/CS8365C 40A/208V Outlets (6) C19/Horizontal NA/JP PDU
		HPE Synergy Composable Infrastructure
3	797740-B21	HPE Synergy 12000 Configure-to-order Frame with 1x Frame Link Module 10x Fans
4	779218-B21	HPE Synergy 20Gb Interconnect Link Module
2	794502-B23	HPE Virtual Connect SE 40Gb F8 Module for Synergy
6	P08477-B21	HPE Virtual Connect SE 16Gb Fibre Channel Module for Synergy
3	798096-B21	HPE 6x 2650W Performance Hot Plug Titanium Plus FIO Power Supply Kit
2	804353-B21	HPE Synergy Composer
2	804937-B21	HPE Image Streamer
3	804938-B21	HPE Synergy Frame Rack Rail Kit
3	804942-B21	HPE Synergy Frame Link Module
1	804943-B21	HPE Synergy Frame 4x Lift Handles
1	859493-B21	Synergy Multi Frame Master1 FIO
1	859494-B22	Synergy Multi Frame Master2 FIO
8	804101-B21	HPE Synergy Interconnect Link 3m Active Optical Cable
2	720199-B21	HPE Blade System c-Class 40G QSFP+ to QSFP+ 3m Direct Attach Copper Cable
2	861412-B21	HPE Synergy Frame Link Module CAT6A 1.2m Cable
1	861413-B21	HPE Synergy Frame Link Module CAT6A 3m Cable
		Virtualized Hosts
3	871940-B21	HPE Synergy 480 Gen10 Configure-to-order Compute Module
3	873381-L21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) FIO Processor Kit
3	873381-B21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) Processor Kit
54	815097-B21	HPE 8GB (1x8GB) Single Rank x8 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit

Qty	Part number	Description
18	815098-B21	HPE 16GB (1x16GB) Single Rank x4 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
6	875478-B21	HPE 1.92TB SATA 6G Mixed Use SFF (2.5in) SC 3yr WTY Digitally Signed Firmware SSD
3	P01367-B1	HPE 96W Smart Storage Battery (up to 20 Devices) with 260mm Cable Kit
3	804424-B21	HPE Smart Array P204i-c SR Gen10 (4 Internal Lanes/1GB Cache) 12G SAS Modular Controller
3	777430-B21	HPE Synergy 3820C 10/20Gb Converged Network Adapter
3	777452-B21	HPE Synergy 3830C 16Gb Fibre Channel Host Bus Adapter (Fibre Channel configs only)
		Worker Nodes
6	871943-B21	HPE Synergy 480 Gen10 6130 2P 64GB-R P204i-c SAS Performance Compute Module
6	873381-L21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) FIO Processor Kit
6	873381-B21	HPE Synergy 480/660 Gen10 Intel Xeon-Gold 6130 (2.1GHz/16-core/125W) Processor Kit
108	815097-B21	HPE 8GB (1x8GB) Single Rank x8 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
36	815098-B21	HPE 16GB (1x16GB) Single Rank x4 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit
6	P01367-B1	HPE 96W Smart Storage Battery (up to 20 Devices) with 260mm Cable Kit
6	804424-B21	HPE Smart Array P204i-c SR Gen10 (4 Internal Lanes/1GB Cache) 12G SAS Modular Controller
6	777430-B21	HPE Synergy 3820C 10/20Gb Converged Network Adapter
3	777452-B21	HPE Synergy 3830C 16Gb Fibre Channel Host Bus Adapter (Fibre Channel configs only)
		HPE 3PAR StoreServ 8440 Storage, Switching, Rack and Power
1	H6Z14B	HPE 3PAR 8440 4N+SW Storage Cent Base
12	K2P89B	HPE 3PAR 8000 1.92TB+SW SFF SSD
1	L7F20AAE	HPE 3PAR All-in S-sys SW Current E-Media
1	P9K08A	HPE 42U 600mmx1075mm G2 Kitted Advanced Shock Rack with Side Panels and Baying
4	AF533A	HPE Intelligent Modular 3Ph 14.4kVA/CS8365C 40A/208V Outlets (6) C19/Horizontal NA/JP PDU
2	C7535A	HPE RJ45 to RJ45 Cat5e Black M/M 7.6ft 1-pack Data Cable
1	Q8B00A	HPE Converged Architecture 750 for Synergy Gen10 without Solution Support Tracking
		iSCSI Connectivity Options
8	N9Z19A	HPE 3PAR 8000 4-pt 10Gb Combo Adapter
		Fibre Channel Connectivity Options
4	H6Z00A	HPE 3PAR StoreServ 8000 4-port 16Gb Fibre Channel Adapter
2	Q0U54A	HPE StoreFabric SN6600B 32Gb 48/24 Fibre Channel Switch
48	P9H32A	HPE B-series 32Gb SFP28 Short Wave 1-pack Transceiver
8	QK734A	HPE Premier Flex LC/LC Multi-mode OM4 2 fiber 5m Cable
48	QK735A	HPE Premier Flex LC/LC Multi-mode OM4 2 fiber 15m Cable
		HPE 5945 FlexFabric Switching
2	JQ075A	HPE FF 5945 2-Slot Switch
2	JH180A	HPE 5930 24p SFP+ and 2p QSFP+ Module
2	JH183A	HPE 5930 8-port QSFP+ Module
4	JH389A	HPE X712 Back (Power Side) to Front (Port Side) Airflow High Volume Fan Tray
4	JC680A	HPE 58x0AF 650W AC Power Supply

Qty	Part number	Description
4	JC680A B2B	INCLUDED: Jumper Cable - NA/JP/TW
2	JG326A	HPE X240 40G QSFP+ QSFP+ 1m DAC Cable
4	JG327A	HPE X240 40G QSFP+ QSFP+ 3m DAC Cable
		Red Hat OpenShift Container Platform
1	R1Z92AAE	Red Hat OpenShift Container Platform for HPE Synergy 1-32 Cores 1yr Subscription 24x7
		Red Hat Enterprise Linux Server
6	J8J36AAE	Red Hat Enterprise Linux Server 2 Sockets 1 Guest 1 Year Subscription 24x7 Support
		Backup Storage (StoreOnce model is a customer choice)
1	BB954A	HPE StoreOnce 3620 24TB System
1	BB982A	HPE StoreOnce Gen4 10/25Gb SFP Network Card
1	BB984A	HPE StoreOnce Gen4 10GbE-T Network Card
1	BB986A	HPE StoreOnce Gen4 16Gb Fibre Channel Network Card

Version history

Project. HPE Reference Configuration for Red Hat OpenShift Container Platform on HPE Synergy and HPE 3PAR StoreServ Storage

Status. Final

Document version	Date	Description of change
1	September 2018	Phase 1 Initial Publication.
2	January 2019	Phase 2 Updates to CA750 Recipe, host automation, install to OCP 3.10, Ansible Tower, Nimble automation, BURA Nimble, BURA 3PAR
3	June 2019	Phase 3 Description and objective of the Business Challenges.
4	December 2019	Phase 4 This update integrates vSphere into the core solution stack, adds GPUs into the solution, consolidates the OCP control plane, adds enhanced security around API authentication build processes, and the registry.

Resources and additional links

Red Hat, redhat.com

Red Hat OpenShift Container Platform 3.11 documentation, https://docs.OpenShift.com/container-platform/3.11/welcome/index.html

HPE Synergy, hpe.com/info/synergy

HPE 3PAR StoreServ Storage, hpe.com/info/3par

HPE FlexFabric 5945 switching,
https://buy.hpe.com/us/en/networking/networking-switches/hpe-flexfabric-5945-switch-series/p/1010907030

HPE OpenShift Solutions on GitHub, https://github.com/hewlettpackard/hpe-solutions-OpenShift

Flexible HPE Converged Systems, https://www.hpe.com/us/en/integrated-systems/converged-architecture.html

Sysdig Solutions, https://sysdig.com/

HPE StoreOnce https://www.hpe.com/in/en/storage/storeonce.html

HPE Recovery Manager Central, https://www.hpe.com/in/en/storage/rmc-backup.html

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

Share now

a00056102enw, Rev. 4, December 2019